David Inman* and Marine Vuillermet

Singular-plural verb stem alternation: uncovering global and local drivers of typological variation

https://doi.org/10.1515/lingty-2023-0040
Received May 24, 2023; accepted June 28, 2024; published online November 22, 2024

Abstract: In this paper, we present a comparative concept for singular-plural verb stem alternation (Sg-Pl alternation) and a systematic global survey of its presence as well as its syntactic and semantic attributes, with a denser sample in the Americas. We show that Sg-Pl alternation is fairly common, occurring in about one in three languages, although its geographic distribution is highly uneven. Fine-grained data collected on the verbs targeted by the phenomenon reveal distinct global and local drivers. General semantic and syntactic properties are fairly consistent at the global level and reveal the relevance of the spatial configuration of the Figure (Motion Event), while at the local level highly specific semantics are shared across neighboring languages. This work will be of interest to both fieldworkers and typologists interested in distinguishing between the various causal factors for the development of typological properties.

Keywords: suppletion; alternation; plurality; verbal number; typology

1 Introduction

The specification of entity number is typically described as a straightforward phenomenon, one which marks plurality on nouns and which may contribute to redundant grammatical marking on other elements, both within the noun phrase (e.g. on determiners, adjectives) and beyond it (e.g. on verbs). Grammatical number is thus a prototypical inflectional category, the semantics of which are centered on the noun but may extend beyond it through agreement (see e.g. Haspelmath 2013 on nominal plurality). Though this description covers a wide range of number

***Corresponding author: David Inman [ˈdeɪvɪd ˈɪnmən],** Department of Comparative Language Science and Institute for the Study of Language Evolution, University of Zurich, Zurich, Switzerland, E-mail: havennah@gmail.com. https://orcid.org/0000-0003-1892-591X
Marine Vuillermet [maˈʁin vɥijɛʁˈmɛ], Department of Comparative Language Science and Institute for the Study of Language Evolution, University of Zurich, Zurich, Switzerland, E-mail: marinevui@yahoo.fr. https://orcid.org/0000-0002-6443-8461

phenomena in the world's languages, we present a typological domain of number marking with a very different set of properties.

The number phenomenon investigated here, which we refer to as SINGULAR-PLURAL VERB STEM ALTERNATION (henceforth SG-PL ALTERNATION), is lexicalized, restricted to the verbal rather than the nominal locus, and excludes all person-based agreement. It refers to verb pairs (or triads)[1] whose alternation is unproductive across the lexicon, and whose components may be very different at the formal level. Semantically, we strictly target alternation triggered by participant number.

A clear example of Sg-Pl alternation can be seen in Yaqui (Uto-Aztecan; Mexico & United States) in example (1).[2] The verb form expressing 'die' changes depending on the number of participants involved in the event: *muuku-* with a singular participant, and *koko-* with a plural participant.

(1) Yaqui (Lindenfeld 1973: 52)
   a. *huʔu čuuʔu* ***muuku****-k*
      this dog **die**(SG)-PFV
      'The dog died.'
   b. *hu-me čuʔu-m* ***koko****-k*
      this-PL dog-PL **die**(PL)-PFV
      'The dogs died.'

In the remainder of the introduction, we review the existing literature on Sg-Pl alternation, and the many (and sometimes contradictory) ways it has been analyzed. In Section 2, we motivate and describe our own typologization.

1.1 History of study

Singular-plural alternation was probably first mentioned by Boas (1911a: 381) in his Tsimshian sketch under the term "irregular plural", but little attention was paid to the phenomenon until the 1980s. As noted by Booker (1982), "even the most comprehensive typology of North American Indian languages (Sherzer 1976) fails to mention number suppletion". It is known by many terms, including "singular versus plural roots" and "number-paired roots" (Kinkade 1981), "number suppletion"

1 All but two languages with Sg-Pl alternation in our sample have at least a bipartite number distinction. In addition, some languages include a third number, like dual or paucal. For the sake of conciseness, we use the term verb "pairs" with triadic verb sets also.

2 For languages in our sample, like Yaqui, the associated glottocode is available in the supplementary material (https://osf.io/amfdn/). For languages not in our sample, the associated glottocode is given in the text at the first mention of the language.

(Booker 1982),[3] "singular-plural root suppletion" (Veselinova 2003), "verbal number pairs" (Veselinova 2006: Chapter 7), and is sometimes reported under the general phenomenon of "verbal number" (Corbett 2000; Frajzyngier 1985; François 2019; Mattiola 2019, 2020; Veselinova 2003, 2013).[4]

Several cross-linguistic works on verbal number have appeared since the early 1980s. Booker (1982), Krasnoukhova (2022), and Mithun (2022) focus on North or South American languages exclusively, while Durie (1986) and Veselinova (2013) are global surveys of the phenomenon. All studies including North American languages emphasize its overwhelming presence there, and Mithun (2022) demonstrates how it is an areal pattern across families in Northern California.[5] Some works also cover more qualitative and/or theoretical topics (Corbett 2000: 243–246; Durie 1986; Frajzyngier 1985; François 2019; Mithun 1988, 2022; Veselinova 2006).

Sg-Pl alternation usually targets both transitive and intransitive verbs (e.g. Mithun 1988, 2022), and less frequently intransitive verbs only (Booker 1982: 20; Durie 1986: 356, 366). The absolutive pattern of the phenomenon has been emphasized since the earliest works: it encodes the number of intransitive subjects and that of transitive objects (Durie 1986: 357; Frajzyngier 1985: 98; Mithun 1988: 214).[6] Durie further observes that this pattern occurs irrespective of the dominant case marking or agreement pattern of the language, and points to a semantic motivation, namely the higher relevance of the "affected argument" (see also Mithun 1988: 214).

The system typically encodes a bipartite opposition: singular versus non-singular or individual versus collective. Booker (1982: 24) and Durie (1986: 356) both note that sometimes languages additionally encode a tripartite opposition, such as

3 Booker's (1982) "number suppletion" covers "only forms characterized by complete root replacement" and thus definitionally excludes reduplication patterns (even if irregularly applied to the verb class) and regular affixational processes where ancient morphology is identified, even if no longer productive.

4 "Verbal number" is sometimes used to refer strictly to loci of marking, contrasting with "nominal number", as in Corbett (2000). Other times it covers the broader semantic domain of number of situations associated with the verb, including both grammatical and lexical morphemes, such as *again*, as in Hofherr and Laca (2012) and Mattiola (2019, 2020). In their work, the grammatical marking of verbal number is called "pluractionality", while others use "pluractionality" to refer to event (rather than participant) plurality, especially in the Africanist tradition (Corbett 2000: 243). For the sake of clarity, we consistently use the transparent terms "event number" and "participant number".

5 Krasnoukhova's (2022) survey shows that verbal number is also present in South America. Her definition encompasses both regular and irregular patterns, whether expressed by reduplication, morphological markers or stem alternation, and expressing event and/or participant number.

6 Booker (1982: 15) actually includes the absolutive pattern in her definition. Durie (1986: 367) mentions an exception noted by Jeanne et al. (1984), but we were unable to access their paper. All languages displaying stem alternation in Krasnoukhova (2022) display an absolutive pattern.

Karok (isolate; United States). Veselinova (2006: 151) reports one language (Maricopa, Cochimi-Yuman; United States) with only tripartite oppositions. The number of verb pairs in a given language varies, from a few to a few dozen (Durie 1986: 356; Krasnoukhova 2022: 642; Veselinova 2006: 153).

In terms of verb semantics, most authors emphasize the salience of motion (spontaneous motion like 'go', and caused motion like 'take'),[7] posture (both spontaneous and caused, like 'lie' and 'lay' respectively) and other frequently-reported semantics which do not fall into a broader category, like 'die' and 'kill' (e.g. Mithun 1988: 232), and sometimes also 'cry' (e.g. Kinkade 1981: 262) and 'sleep' (e.g. Veselinova 2013). Some languages display such an alternation in some adjective roots, often 'small' and 'big' (Booker 1982; Durie 1986: 361; François 2019: 349, 351; Krasnoukhova 2022: 651; Mithun 1988: 218; Veselinova 2006: 154).

1.2 Different approaches

While there has been agreement on some aspects of Sg-Pl alternation – its prevalence in North America, absolutive patterning, the greater frequency of intransitive verbs, and the semantic skew of alternating verbs – there have also been considerable differences in how the phenomenon has been approached. The very definition of Sg-Pl alternation has varied substantially across researchers.

All authors consider verb pairs which display unexpected formal differences (often referred to as "proper suppletion"), and most also include verb pairs where the alternation is encoded by a non productive (plural) morpheme, irregular ablaut or reduplication included. These non-suppletive pairs are irregular in that their plural morphology is restricted to a closed set of verbs, much like the suppletive pairs. Veselinova (2006: 158–169) considers the difference between alternating verbs which form exceptions to otherwise regular agreement paradigms and verbs which are not part of a paradigm, concluding that these are two ends of a continuum (p. 173). This distinction between paradigmatic ("suppleting") and non-paradigmatic ("non-suppleting") alternation is not made by all authors.

Perhaps the most significant discrepancy is the semantics of "verbal number" (see footnote 4). While some authors distinguish between participant and event number (Corbett 2000; Durie 1986; Krasnoukhova 2022; Mithun 2022), others consider them jointly (Frajzyngier 1985: 98; Mattiola 2019, 2020; Mithun 1988; Veselinova 2006,

7 Mithun (1988: 232, 2022) also mentions handling verbs, but the list she gives ('put', 'take', 'pick up', 'drop', 'throw') could alternatively be analyzed as caused motion verbs. Booker (1982) calls "locative verbs" what we call "posture", "location" ("existence" in her terminology) and "motion verbs". Her "non-locative verbs" include among others the frequently-reported semantics of 'die', 'sleep' and 'cry'.

2013). Only a few individual language studies like Heath (1980) or François (2019) clearly set apart the two semantic domains. In others it is more complex, as in Canela-Krahô (Nuclear-Macro-Je; Brazil), where Miranda (2014: 134) separates the two, although he acknowledges that plurality of arguments is *also* considered by the speakers as indicating a plurality of events – e.g. killing several people can be considered as a single killing event or as several distinct killing events (Maxwell Miranda, pers. comm. December 2022, in the same spirit as Frajzyngier 1985: 98). In still other studies, sources disagree on the semantic interpretation of the verb pairs within a single language, as is the case for Everett (2006: 348–355) and Storto (2014: 409–410) on Karitiâna (Tupian; Brazil). This is one reason why Corbett (2000: 248–249) and Krasnoukhova (2022: 641–642) classify the phenomenon into three subtypes: event plurality only, argument plurality only, and a last mixed type. Interestingly, Müller (2014) and Krasnoukhova (2022: 643) note that "proper suppletion" necessarily expresses the number of participants, sometimes along with that of events.

Another theoretical disagreement concerns the syntactic or lexico-semantic nature of the phenomenon. Several studies, starting with Boas (1911b: 381), insist that the phenomenon is distinct from agreement (i.e. echoing the nominal number marked on NPs), and is rather associated with a lexical choice based on the overall semantics of the clause. The three main reasons for an analysis in terms of semantic selection (rather than syntactic redundancy) are:

1. In many cases, no (explicit) NPs are present to agree with the verbal number (e.g. Frajzyngier 1985: 96–97; Mithun 1988: 212);
2. Syntactic incongruence occurs in several languages (e.g. I 1SG-walk:DUAL with him vs. one of us 1NON.SG-walk:SG in Navajo (Athabaskan; United States) as in Durie 1986: 358; see also Mithun 1988: 215);
3. Sg-Pl alternation is part of a larger phenomenon targeting not just participant number but also specific participant shapes (Mithun 1988: 214, 2022).

Additionally, François (2019: 357) describes how Sg-Pl alternation in Hiw [hiww1237] (Austronesian; Vanuatu) follows semantic rules that are specific to verbal number (SG-PL), and only partially coincide with nominal number (SG-NON.SG). Note however that Veselinova (2013) shows that analyses in terms of lexical selection or agreement might both be present within a given family, e.g. in Athabaskan, where the syntactic incongruence described in Navajo (see (2) above) is impossible in Slavey.

For any typological survey, the consistent coding of variables across potentially very different languages is fundamental, and it requires strict and clear definitions. We now describe how we divide the conceptual space, i.e. the "comparative concepts" (Haspelmath 2010) we use in our survey.

2 Definitions

We have called our central comparative concept singular-plural verb stem alternation and defined it as (1) LEXICALLY SPECIFIC SETS of verb stems, (2) whose distribution can be determined by the NUMBER OF AN ARGUMENT.

We chose the term singular-plural verb STEM alternation because formally our definition encompasses two types of alternation (discussed in detail below): root suppletion, and lexically limited regular pattern. Wari' (Chapacuran; Brazil) displays both types: it has 14 suppletive pairs, and 36 lexically limited pairs, either obtained by reduplication in the form of *CV(CV)-* or reduplicative infixation in the form of *-rV(CV)-*, as illustrated in Table 1. We include both of these alternation types because, to our knowledge, previous studies have not discovered cross-linguistic differences according to this formal parameter.

Semantically, our comparative concept of Sg-Pl alternation excludes event number, and only targets participant number. This restriction is made primarily because participant and event number describe different semantic domains and are conceptually separable. In fact, there are languages which encode the two separately. This is the case in Seri (isolate; Mexico), where different morphological processes encode event number (repetition) and participant number, giving rise to a four-way distinction for most verb forms. The full morphology governing these forms is quite complex, but event and participant number can be marked by affixation processes alone (as in 'talk about', Table 2) or by affixation and suppletion (as in 'go to', Table 3).[8] See also Hiw for a very clear four-way distinction (François 2019: 353–354).

Participant number can be straightforwardly defined (singular vs. multiple entities), while event number encompasses a variety of concepts. The latter can refer

Table 1: Two types of Sg-Pl alternation in Wari' (Birchall et al. to appear).

Root suppletion			Regular pattern		
	SG	**PL**		**SG**	**PL**
'stand'	*xat*	*koko'*	'hang'	*wa*	*wa**rawa***
'run'	*mao*	*mama'*	'walk'	*xut*	*xu**ruxu**t*
'die'	*mi'*	*kono'*	'kill'	*topa'*	*topa**rapa**'*

8 Note that in Table 2, suffixes encode both participant and event number. In Table 3, suffixes again encode both participant and event number, but root alternation encodes participant number only.

Table 2: The verb 'talk about' in Seri (Marlett 1981: 96, 224).

	Talk about (single P argument)	Talk about (multiple P arguments)
One event	*-šaχ^{w}*	*-šaχ^{w}**-t***
Multiple events	*-šaχ^{w}**-tim***	*-šaχ^{w}**-toɬka***

Table 3: The verb 'go to' in Seri (Marlett 1981: 96, 224).

	Go to (single S argument)	Go to (multiple S arguments)
One event	***-ya:i***	***-oši:t***
Multiple events	***-ya:i**-tim*	***-oši:t**-am*

to, at least, various repetitive aspects[9] and habitualness. These types of event plurality are distinguished according to the interval between repetitions, and it is difficult to differentiate between repetitive and habitual repetition in a cross-linguistically consistent way. Both types of event repetition are problematic, as many potential verb pairs will have to be considered as displaying singular-plural (event) alternation, such as English *step* ~ *walk*, *stroke* ~ *pet*, and *think* ~ *ponder*.

Nevertheless, in some languages, the same process can be used to indicate event number in some verb roots and participant number in others. In Choctaw (Muskogean; United States), the verb 'put down', for instance, suppletes for participant number, while 'hit' suppletes for event number (Heath 1980: 2–3, 21–22).

When a source explicitly specified that the alternation encodes argument (and not just event) number, we have followed the author. In ambiguous cases, we looked for examples which clearly show participant number, to the exclusion of event number. For intransitive verbs, this is relatively straightforward (i.e. multiple entities performing the action, rather than one entity repeating it), but for transitive verbs we looked for one of the following contexts:

– A single A(gent) repeats the action on multiple P(atient)s of the same type.
  Ex: One person shoots multiple targets
– Multiple A(gent)s perform the action.
  Ex: Multiple people shoot (perhaps the same target)

In some languages, some stem alternations can be triggered by either participant or event number. Since participant alone is sufficient, such stems are included.

9 In some languages, there are multiple repetitive aspects, as in Nuu-chahn-nulth (Wakashan; Canada), where the regularity of repetition is encoded as lexical aspect (Nakayama 2001: 27–28). The form indicating irregular repetition is occasionally lexically irregular.

Our definition also entails that we exclude cases where additional semantic factors distinguish the verbs. As defined by Mithun (2022), this means that alternating verbs must be "basic-level terms" and not in a synchronic hypernym/hyponym relationship. An example of a verbal pair that fails to meet this criterion is English *kill*/*massacre*. The verb *massacre* necessarily involves a plural object, but it is a hyponym of *kill*, i.e. not a basic-level term. More than just a plural object, *massacre* also provides information about brutality, intentionality, a short span of time, and other semantic properties. Consequently, *kill* can always felicitously replace *massacre* with a loss of specificity (as in (2)), but *massacre* cannot always felicitously replace a *kill* with a plural object (as in the semantically odd (b) sentences in (3) and (4)).

(2) a. *The invaders killed the villagers.*
b. *The invaders massacred the villagers.*

(3) a. *The traitors were sentenced to be killed by hanging.*
b. *# The traitors were sentenced to be massacred by hanging.*

(4) a. *I inadvertently killed 100 insects with the windshield over the month I rented the car.*
b. *# I inadvertently massacred 100 insects with the windshield over the month I rented the car.*

Another type of alternation that we exclude definitionally is suppletion that encodes person along with number, as in Malinaltepec Me'phaa (Otomanguean; Mexico) (see Table 4). Though the paradigm shares the same stem amongst all plural subjects, the forms for singular subjects change based on person. Had the stem remained consistent in the singular, then it would have counted as a form of Sg-Pl alternation.

Finally, our definition does not specify the type of alternation verbal stems undergo, and thus encompasses two formally distinct types of alternation: root suppletion, and lexically limited plural morphology. ROOT SUPPLETION (henceforth SUPPLETION) is when the verb roots are fully suppletive, as with *-ya:i* ~ *-oši:t* 'go to' in Seri (Table 3), or where no pattern is systematic across multiple verb pairs.

The other type of alternation, which we call LEXICALLY LIMITED REGULAR PATTERN (henceforth REGULAR PATTERN), is when verb stems alternate by segmentable but

Table 4: The verb 'eat' in Malinaltepec Me'phaa (Suárez 1983: 160).

	SG	PL
1	*-kho²*	*-pho²*
2	*-'co²*	*-pho²*
3	*-kho²*	*-pho²*

Table 5: Infixing <*l(V)*> ~ <*n(V)*> plurality in Kunama[a] (Böhm 1984: 28–29; Tucker and Bryan 1966: 340).

	SG	PL
'arrive'	*i*	*i**l**i*
'come'	*(i)o*	*(i)**l**o*
'enter'	*u*	*u**l**u*
'run'	*egi*	*i**l**agi*
'dwell, stay'	*uta*	*u**l**uta*
'wash'	*ufe*	*u**l**ufe*
'steal'	*una*	*u**n**una*
'weep'	*imbi*	*u**n**imbi*

[a]There is a copy vowel for the <*l(V)*> infix in every word except 'come' and 'run'. The plurals of 'run' and 'weep' look close to suppletion. However, not knowing the full morphophonology of the language, we believe it is plausible that the relationships between the singular and plural forms are phonologically straightforward, especially given that the author does not explicitly mention the irregularity.

unproductive morphology, including reduplication, in at least two verb pairs (if in only one, it is root suppletion).

A regular pattern must be lexically limited, and so definitionally excludes semantically restricted morphology. For instance, in Magdalena Peñasco Mixtec (Otomanguean; Mexico), the morpheme *koo* marks a plural subject only and for all motion verbs (de Hollenbach 2013: 48).[10]

Kunama (isolate; Eritrea, Ethiopia, and Sudan) displays a regular pattern with eight verbs that have an unproductive <*l(V)*> or <*nV*> infix marking plurality (Table 5).[11] These verbs do not exhaustively cover any semantic domain: though the first four verbs are motion verbs, not all motion verbs display the infixation.

In most languages, regular patterns are expressed by plural affixation (including reduplication) to the singular verb form, as is the case in Kunama. However, in one language of the sample, regular alternation is expressed via affixation to the plural form: Daga (Dagan; Papua New Guinea) has a singulative prefix *wa-* for an unspecified number of verb roots (Murane 1974: 11–12), as in (5).

10 This semantic restriction appears to be motivated by the historical origin of *koo* from *koyo* 'pour', or metaphorically 'movement in a group, pouring out' (de Hollenbach 2013: 151).

11 There is a copy vowel for the <*l(V)*> infix in every word except 'come' and 'run'. The plurals of 'run' and 'weep' look close to suppletion. However, not knowing the full morphophonology of the language, we believe it is plausible that the relationships between the singular and plural forms are phonologically straightforward, especially given that the author does not explicitly mention the irregularity.

(5) Daga (Murane 1974: 41)
   a. ***wa****-dum* ‘split (SG object)’
   b. *dum* ‘split (PL object)’
   c. ***wa****-boR* ‘close (SG object)’
   d. *boR* ‘close (PL object)’

Note that the difference between regular pattern and suppletion is sometimes unclear. In Western Keres (Keresan, United States), many verb forms seem regular, but are not in their entirety: *-uːbəN* ~ ***-âaʔ****abəN* ~ ***-jɛ̀ɛ̀****bəN* ‘enter’ and *-jaʔaċíN* ~ ***-âaʔ****áċiN* ~ ***-jée****dyuB* ‘arrive’ share the same dual prefix, but not quite the same plural prefix. Furthermore, the root for ‘enter’ is consistent across all numbers while that for ‘arrive’ is suppletive in the plural (Miller 1965: 59). We followed Miller’s analysis and considered such cases suppletive.

There is a very particular kind of edge case with irregular morphology for participant plurality: a language may have pluralizing morphology for *all* verbs, but the form that it takes is unpredictable and lexically conditioned. Nisga’a (Tsimshianic; Canada) exemplifies such a system, as described in great diachronic detail in Tarpent (1983). Its plural system developed in stages, with later morphology modifying or adding to existing plural (or distributive) morphology, resulting in a contemporary system with at least eight strategies for marking a plural absolutive argument: *lV-* prefixing, *tx̌a·* - prefixing, *Cix-* prefixing, *qa-* prefixing, full stem reduplication, irregular reduplication (i.e. with phonologically irregular onsets or codas), partial reduplication, and full reduplication. Additionally, there are some verbs that undergo root suppletion to mark plurality. Every verb marks the plural of its absolutive argument using one of these strategies, but there is no way to predict which strategy a verb will use. We see two ways to interpret systems like Nisga’a’s:

1. Every verb has unpredictable verbal morphology, and thus every verb meets our definition of Sg-Pl alternation by regular pattern. These languages have entire lexicons of alternating verb pairs.
2. These systems encode plurality via conjugation paradigms, similar to person and tense systems in many European languages. There is a single plural paradigm which applies to every word, but the form is lexically conditioned by the (plural) inflectional class to which the verb belongs. There is thus predictable (but abstract) morphology for plurality with lexically-dependent allomorphy.

Languages like Nisga’a represent a point at which lexical irregularity and consistent grammatical marking meet. We have selected the second interpretation for comparative reasons. Such plural alternations encompass the entire verbal lexicon, while in most other languages they are limited to a small, closed set.

We think it is best to compare across closed sets, and two observations seem to confirm this decision. First, only one of sixteen languages with unpredictable plural

marking that encompasses the entire lexicon lacks Sg-Pl suppletive roots (Salinan, isolate; United States).[12] Second, as we shall show in our results (Section 4), the semantics of the verb pairs that undergo Sg-Pl alternation are far from random, an observation that extends to languages with unpredictable verbal plurals. This generalization would be obscured by including the entire lexicon.

For similar reasons, we have considered only underived verb pairs and systematically excluded derived ones, such as 'lay X' if derived from 'lie'. This decision avoids artificial inflation of verb pairs in languages with extensive and productive morphology.

Finally, in addition to (lexical) stems, we investigated Sg-Pl alternation in verbal morphology. Of course, such morphemes encode not just number distinctions but also a grammatical category such as causation or aspect. For instance, Tohono O'odham (Uto-Aztecan; United States; Saxton et al. 1983: 41) has a transitivizing suffix pair *-wu* (SG) versus *-shulig* (PL) which alternate according to the subject's number.

3 Sample and methodology

The language sample for this study was determined by a larger typological project (the Out of Asia SNSF Sinergia project) targeting the Americas. We thus have an oversampling of American languages (220) compared to languages from the rest of the world (105), for a total of 325 languages (see Figure 1).[13] Language names, glottocodes, genetic affiliation, and geographic coordinates follow *Glottolog* (Hammarström et al. 2021). Within each macroarea (as defined in Hammarström and Donohue 2014), our language sample maximizes phylogenetic diversity (and isolates over families), while trying to cover as much geographic space as possible. As a result, it is almost only in the Americas that a language family has several representatives. We have also avoided LOL languages (Literate, Official languages with Lots of speakers, see Dahl 2015), since their atypical development would obscure the baseline comparison with the indigenous American languages.[14]

12 The other fifteen languages are: Hopi, Maricopa, Seri, Tohono O'odham, Western Keres, Zuni, Kashaya, Washo, Yurok, Upper Chehalis, Southern Coastal Tsimshian, Nisga'a, Koasati, Kunama, and Imonda. All are spoken in North America except Imonda, spoken in Papunesia. Creek and Laal almost make it into this category, with irregular plural morphology for over 100 verbs.

13 Maps and figures in this paper were created using the sf (Pebesma 2018; Pebesma and Bivand 2023) and ggplot2 (Wickham 2016) R packages, Natural Earth (http://www.naturalearthdata.com), and the D3.js package (https://d3js.org/).

14 Behavioral scientists have used the concept of "Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies" (Henrich et al. 2010: 61) to criticize earlier generalizations about human psychology and behavior mainly based on this unrepresentative group. Dahl (2015) extends

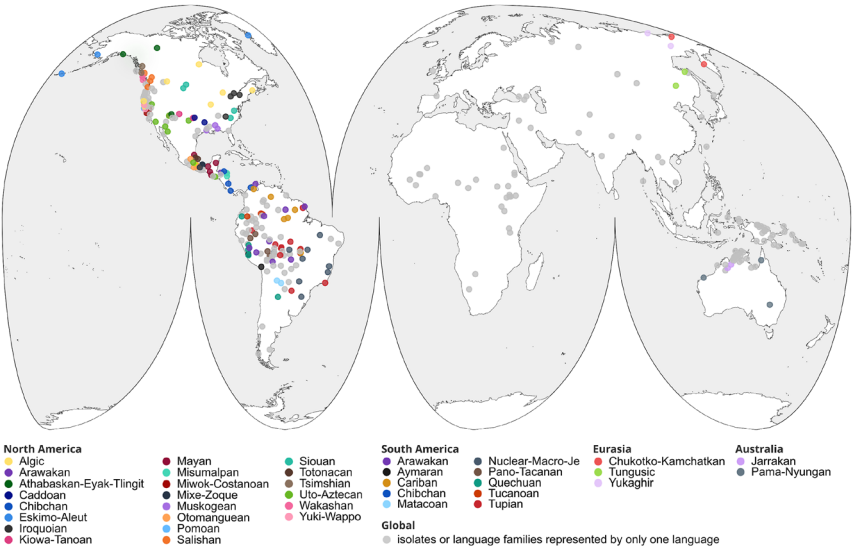

Figure 1: Language sample for the study.

The sampling bias means that our results are most indicative for the Americas. Our results for Africa, Eurasia, Papunesia, and Australia may under- or over-estimate the prevalence of Sg-Pl alternation, depending on the profile of very large families in those macroareas, most of which have only a single representative in our sample.

Our data are organized in two tables, both available in the online supplementary materials (see the Data availability statement). The main one includes 21 typological questions with limited possible answers, followed by associated detailed remarks and sources. These cover the morphological type and number of Sg-Pl pairs (see Section 4.1), their syntactic properties and the triggering argument (Section 4.2), the number distinction they encode (Section 4.3), and their lexical semantic category (Section 4.4), alongside other properties which did not reveal any larger pattern and are not discussed here, such as the use of plural forms to indicate iterativity, the presence of unaccusative and unergative intransitives, or the additional presence of suppleting adjectives.

The other table provides the forms of each alternating verb pair, their free translation, and alternation type (suppletion or regular pattern, and bi- or tripartite pattern). Additional columns for semantic categorization were added based on a semantic gestalt that arose from the table. These data inform the more detailed parts

this concept to linguistics, and shows, for instance, that languages with “Lots of speakers” correlate with a high level of contact and many second language speakers, properties claimed to be correlated with low morphological complexity and large phonological inventories.

of our semantic analysis (see Section 4.4) and the local spread of the phenomenon (Section 5.2).

Automatic error and warning checks were written in R. Scripts checked for definitional inconsistencies (e.g. if the language has intransitive alternating pairs, they must be unaccusative, unergative, or both), state combinations we assumed a priori to be rare, and consistency between verb counts in the two spreadsheets. We also went through Mattiola's (2019) list of stem alternations to verify that the languages present in both samples had been coded consistently.[15]

Most of our data came from grammars and dictionaries. If available, dictionaries were systematically checked for alternating stem pairs, manually if not automatically searchable. For complicated cases or cases of poor documentation, we reached out to the target language specialist(s) (see Acknowledgments).

4 Global results

Our survey has uncovered two types of patterns present in Sg-Pl alternation. This section presents the typical syntactic and semantic properties consistent across macroareas, while Section 5 details the local distribution specific to smaller areas, which suggest particular cases of language contact.

4.1 Distribution

Sg-Pl stem alternation is fairly frequent, occurring in a third of our sample (109/325 languages). Its distribution across macroareas is highly uneven, with North America greatly overrepresented (Figure 2). The raw counts and frequencies of Sg-Pl alternation are given in Table 6, while Table 7 presents a stochastic subsampling method, described below.

As discussed, our sample is heavily skewed toward the Americas, which happen to occupy opposite ends of the distribution. We have tested and corrected for any resulting biases in two independent ways: first, by rebalancing the raw frequencies so that macroareas are weighted evenly (presented in the last row of Table 6), as described in this paragraph, and second, by performing random subsampling (Table 7), as described further below. Rebalancing the raw frequencies in Table 6 by macroarea (i.e. weighting each macroarea by 1/6) gives only a slightly lower global

15 We had very few disagreements on the semantic analysis of the verb pairs, though in some cases we considered phenomena analyzed by Mattiola as expressing event number to rather be expressing participant number.

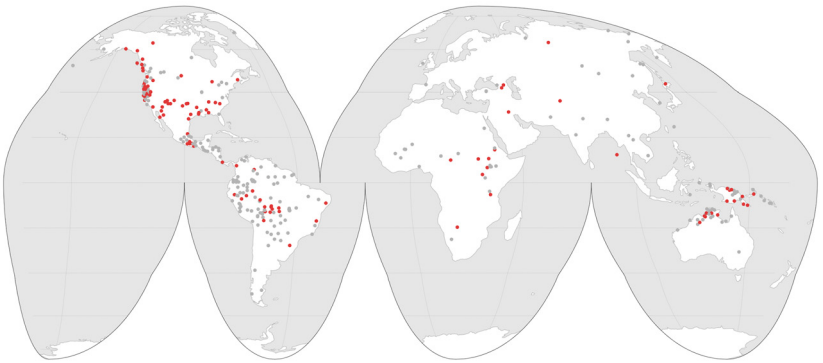

Figure 2: Sg-Pl verb stem alternation.

Table 6: Presence of Sg-Pl alternation.

Macroarea	Frequency	Count
North America	52 %	60/115
Papunesia	36 %	9/25
Africa	32 %	8/25
Australia	24 %	6/25
Eurasia	23 %	7/30
South America	18 %	19/105
Global	33 %	109/325
Rebalanced	*31 %*	
Without North America	*27 %*	

Table 7: Sg-Pl alternation prevalence over 250 subsamples.

Macroarea	Mean	Standard deviation
North America	53 %	12 %
Papunesia	36 %	8 %
Africa	31 %	7 %
Australia	24 %	7 %
Eurasia	25 %	8 %
South America	18 %	10 %
Global	31 %	4 %

average of 31 %. Rebalancing without the outlier of North America gives a still lower average of 27 %. If Central America (defined as north of South America and south of the Chihuahuan Desert) is separated from North America, 60 % (52/86) of the remaining languages of North America have at least one verb pair, while the languages of Central America display a much more globally typical 23 % (7/30). In other words, Sg-Pl alternation is more than twice as frequent in North America than in the rest of the world, a matter we will return to in the discussion.

A different balancing method (Table 7) was implemented, following concerns from a reviewer about remaining macroareal biases. We generated 250 independent, random subsamples of 15 unrelated languages per macroarea (90 languages in total), guaranteeing that there was no phylogenetic or geographic bias in a given sample. From these samples, we calculated the mean and standard deviation for the frequency of the presence of Sg-Pl alternation in each macroarea. As seen by comparing Tables 6 and 7, the resulting means are nearly identical to the overall frequencies, though standard deviations vary by macroarea. The global standard deviation is however notably low. This procedure confirms that the American skew of our sample is not biasing macroareal differences. These frequencies are similar to the figures in Veselinova (2006: 156), except that we find a much greater presence of Sg-Pl alternation in Papunesia.

As will be detailed in Section 5.1, the consistency of Sg-Pl alternation within a family is somewhat mixed. In some families, all or almost all languages display the phenomenon (e.g. the Athabaskan-Tlingit-Eyak, Uto-Aztecan, and Otomanguean families), in others only half of the languages display it (e.g. the Pano-Tacanan family), in still others it is even less (e.g. the Algic, Salishan, and Tupian families).

Within a language, typically only a handful of verbs alternate. Figures 3 and 4 group languages according to the number of verb pairs that alternate by suppletion and by regular pattern, respectively. Recall that we have excluded derived verb pairs.

Root suppletion is the most common way of forming Sg-Pl verbal pairs (present in 104/109 languages in our sample). The mode number of suppletive pairs is 6 and the mean is 9, with half the languages having 7 or fewer suppletive pairs. The mean is substantially higher than the mode because of a long tail distribution, with a handful of North American (and one Papunesian) languages having very large numbers of verb pairs: at the extreme tail, 34 in Hopi (Uto-Aztecan), 46 in Comanche (Uto-Aztecan), and 61 in Western Keres (Keresan). Sg-Pl alternation by a regular pattern has a similar distribution, although this strategy is much less common overall (32/109 languages).[16]

16 Alternation by regular pattern is most typically accompanied by at least some suppleting verb pairs. Only five languages in our sample exhibit regular patterns only, with no root suppletion: Mullukmulluk, Murriny Patha, and Wubuy in northern Australia, Kunama in Africa, and Ottawa in North America.

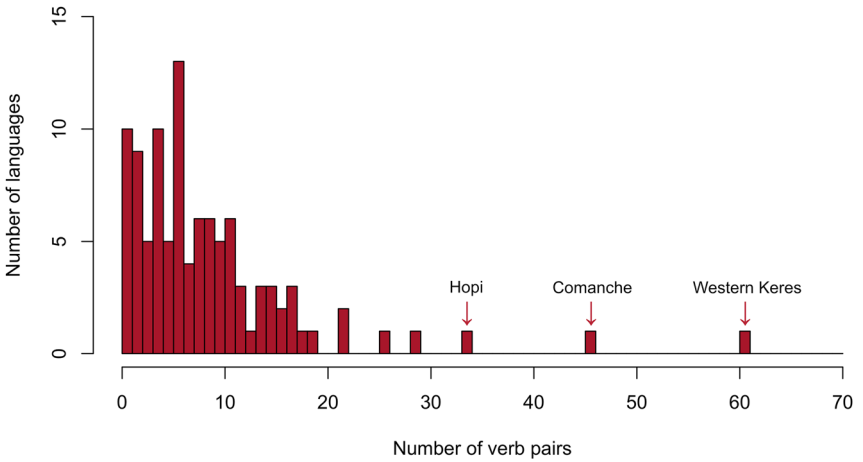

Figure 3: Root suppletion counts by language (*N* = 103).

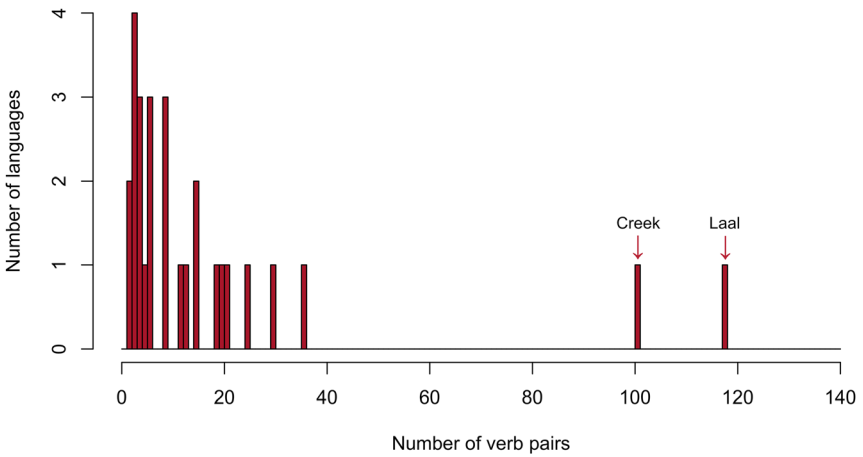

Figure 4: Regular pattern counts by language (*N* = 28).

The mode number is 3, and the mean is 18, with half the languages having 9 or fewer verbs following a regular pattern. The mean is much higher than the mode once again because of a small number of outliers, in particular Creek (Muskogean) and Laal (isolate) with over one hundred pairs.

As for alternation in verbal morphology, of the 33 languages in our sample which display it, six lack Sg-Pl alternation in main verbs, a pattern we expected to be much rarer. Four of these languages are in or near the Andes (Huallaga Huánuco Quechua,

North Junín Quechua, Urarina, and Wichí Lhamtés Nocten), possibly indicating a regional historical process. The other two are far apart (Kusunda in Eurasia and Northwest Maidu in North America).

Our sample confirms the preponderance of Sg-Pl stem alternation in North America. However, it is still common in the rest of the world, occurring in about one in four languages in our sample outside of North America. As expected for a typological feature which is definitionally lexically limited, it typically occurs in a handful of verbs in a language, and significantly, suppletion is much more common than limited regular patterns.

4.2 Syntactic profile

Previous researchers (e.g. Booker 1982: 20) have noted that Sg-Pl alternation targets either only intransitive verbs or both intransitive and transitive ones. We confirm this tendency, with 33/109 languages having only intransitive alternating pairs, and 70/109 having both intransitive and transitive pairs. However, five languages in our sample show alternation only for transitives.[17] Of these, only the extinct and little-documented language Esselen (isolate; United States) is from North America. The remaining languages are in Papua New Guinea (Daga, Nimboran, and Yau-Nungon) and the Andaman Islands (Akabea), regions where Sg-Pl alternation has been scarcely investigated. The cross-linguistic tendency for intransitive verbs may not hold for the whole of Papunesia, where the languages in our sample never display intransitive-only alternation, but either transitive-only or both.

We confirm the overwhelming tendency for Sg-Pl alternation to be triggered by S and P roles (Corbett 2000: 253; Frajzyngier 1985: 98; Mithun 1988: 214). In 62 out of 75 languages with transitive alternating pairs, only the P role is targeted. However, in five languages, the A role is targeted: Laal (isolate; Africa), Murriny Patha (Southern Daly; Australia), Akabea (Great Andamanese; India), Seri (isolate; United States), Malinaltepec Me'phaa (Otomanguean, Mexico); and in six languages, some transitive pairs target the A role, and others the P role: Ket (Yeniseian; Russia), Coahuilteco (isolate; United States), Hopi (Uto-Aztecan; United States), Imonda (Border; Papua New Guinea), Jaminjung-Ngaliwurru (Mirndi; Australia), Wari' (Chapacuran; Brazil).[18] No language in our sample displays verb pairs which could alternatively express A or P plurality.[19]

17 One language was uncategorizable due to insufficient data.

18 Nakagawa (2022: 594) thoroughly explains why he does not consider Sg-Pl alternation in Hokkaido Ainu (Ainu; Japan) to be accusatively aligned, contra Shibatani (1999: 50–54) and Corbett (2000: 253). For another example of an accusatively aligned Sg-Pl alternation and its historical development, see François (2019: 350).

19 For two languages, there was insufficient documentation.

Verbal morphology follows a similar absolutive pattern. The associated motion suffix pair *-m(m)ed* (SG) versus *-op(po)* (PL) 'GO.TO.DO' in Tohono O'odham may co-alternate with a lexical verb pair (here, *me'a* ~ *kokthai* 'kill'), yielding the four possible combinations illustrated in (6). The alternating verb pair targets the number of P, and the alternating suffix pair seems to target A. However, the suffix actually targets the S role of an associated intransitive Motion Event, the expression of which is grammaticalized.

(6) Sg-Pl verbal morphology in Tohono O'odham (Saxton et al. 1983: 41)

a. *me'a-**med***
kill.SG-**GO.TO.SG**
'go (S.SG) to kill (P.SG)'

b. *me'-**op***
kill.SG-**GO.TO.PL**
'go (S.PL) to kill (P.SG)'

c. *kokthai-**med***
kill.PL-**GO.TO.SG**
'go (S.SG) to kill (P.PL)'

d. *kokthai-**op***
kill.PL-**GO.TO.PL**
'go (S.PL) to kill (P.PL)'

4.3 Number distinctions

Bipartite number distinctions are overwhelmingly more present than tripartite ones, with 97/109 languages showing bipartite only distinctions, 14 showing both bi- and tripartite distinctions, and two showing tripartite distinctions only. Among the bipartite number distinctions, the semantics is overwhelmingly Singular-Plural (see Table 8).[20]

The Singular-Plural number distinction refers to a strict singular versus non-singular opposition, Individual-Collective opposes an individual acting alone versus a group acting collectively, and Paucal-Many a small versus large number. Tripartite distinctions generally add a dual rather than a paucal as an intermediate step between singular and plural (see Appendix A for the detailed distribution of tripartite patterns).

Individual-Collective number systems being less frequent in the literature, we illustrate with two examples from a single paragraph in a Nuu-chah-nulth

20 Note that some languages have different types of number contrasts for different verbs, hence the total is above 109.

Table 8: Number distinctions.

System	Number distinction	Number of languages
Bipartite	Singular-Plural	100
	Individual-Collective	8
	Paucal-Many	7
Tripartite	Singular-Dual-Plural	15
	Singular-Paucal-Plural	1

(Wakashan; Canada) text. In (7), the geese fly up from a beach as a group and the collective form for 'fly' *huʔak* is thus used. In (8), the geese have been blinded by fires and no longer fly collectively, and the individual form *mataa* is used despite a plural subject. That is, collectivity, rather than plurality, determines the choice of verb form.

(7) Collective flying in Nuu-chah-nulth (Sapir and Swadesh 1955: 13)
ʔuunuuƛ ʔiiḥ ʔuusaʕaƛ ***huk***^{w}*isaʔaƛquu ʔuyi ʔuunuuƛ ʔiiḥ* ***ʔaayuu huk***^{w}***iƛ ḥuqsim.***
ʔuunuuƛ ʔiiḥ ʔuusa-ʕaƛ ***hu-****k*^{w}*is-saƛ=!aƛ=quu*[21]
because big/very loud-make.a.sound **fly.**COLL-up-on.beach.PFV=NOW=POSSIBLE.3
ʔuyi ʔuunuuƛ ʔiiḥ ***ʔaya-L.uu*** ***hu-****k*^{w}*iƛ* ***ḥuqsim.***
at.a.time because big/very **many-as.one** **fly.**COLL-PFV **goose**
'There were so many **geese flying** up from the beach that it made a loud sound.'

(8) Individual flying in Nuu-chah-nulth (Sapir and Swadesh 1955: 13)
ʔicʔinksawiʔaƛ maamaati ḥisiɬiičiʔaaḥʔaƛ saye?ii ***mat****aa taakʕaʔataƛ ʔuunuuƛ ʔicʔinksawiʔaƛqa.*
ʔink-R<c>L.sawiƛ=!aƛ maamaati ḥi-siła-iičiƛ=!aaḥ=!aƛ saya=ʔii
fire-at.eyes.PFV=NOW bird unable.to-do-PFV-IRREALIS=NOW far=go.to
mat*-aa tak-ʕaaʔatu-LS=!aƛ ʔuunuuƛ*
fly.INDV-CONT straight-downward-ONGOING=NOW because
ʔink-R<c>L.sawiƛ=!aƛ=qa
fire-at.eyes.PFV=NOW=EMBED
'The birds, blinded by the fires, were unable to **fly** far.'

The Nuu-chah-nulth verbs *mataa* and *huʔak* are thus basic-level terms, unlike the English verbs *fly* and *flock*, which are in a hypernym/hyponym relationship. While *fly* is agnostic about the individuation or collectivity of its (possibly plural) subject, *flock* necessitates a collective subject.

21 Nuu-chah-nulth has a rich, partially lexicalized system of aspect: *mataa* is the continuative form of 'fly.SG', and *matšiƛ* its perfective; *huʔak* is the continuative of 'fly.PL', and *huk*^{w}*iƛ* its perfective.

Such a determination is dependent on access to large corpora, which are not available for many languages. Even when Sg-Pl alternation is discussed in a grammar, thorough semantic discussion is often lacking. Following Corbett (2000: 250) and the observations of Boas (1911a: 381) about the importance of collectivity in many American languages, the individual-collective opposition could be more frequent than our results suggest.

The presence of multiple number distinctions within a single language is rare (18/109), and most typically involves the addition of a tripartite distinction to an otherwise bipartite system (14/18 mixed systems, presented in more detail in Appendix A). Only a few languages cumulate different bipartite systems without any tripartite ones, like Yuracaré (isolate; Bolivia). Van Gijn (2006: 191–192) shows how the verb pairs for 'go', 'arrive', and 'go across' strictly follow a Singular-Plural semantic distribution, while other verb pairs like 'fall' follow a Paucal-Many distribution, and are grammatical with low numerals.

4.4 Lexical semantics

The lexical semantics of verb pairs follows previous observations. We report here only on the figures for suppletion for the 103 languages where we have lists of suppletive pairs, since looking at root suppletion only or at all stem alternations makes little difference to the overall picture.[22] We distributed verb pairs into five semantic categories: Motion, Posture, Location, Frequent Verbs, and All Others, as illustrated in Table 9. Motion, which includes both spontaneous and caused motion verbs (like 'go' and 'take'), is the most predominant semantics, both in terms of number of languages (86/103) and total number of suppletive pairs (379/928).

The Posture category, which includes both spontaneous posture (like 'lie') and caused posture verbs (like 'lay'), is the second most predominant semantics, again

Table 9: Attested verb semantics of suppletive alternations by languages and by tokens.

Categories	Languages (*N* = 103)	Tokens (*N* = 928)
Motion (e.g. go, take)	86	379
Posture (e.g. sit, lay down)	67	190
Location (e.g. exist, be there)	35	50
Frequent Verbs (see Table 10)	52	99
All Others	55	210

22 For example, the Motion category occurs in 89 languages in all stem alternations versus in 86 in root suppletion only, and Posture in 71 vs. 67 languages.

both in terms of number of languages (67/103) and tokens (190/928). Note that verbs like 'put X (lying)' or 'lay X' express both caused motion and caused posture, and were sometimes not easy to categorize. We followed the cues given in the translation in the sources, unless the verb form indicated the opposite analysis.

The Location category includes location verbs (like 'be there' or 'exist') and caused location verbs (like 'hold' or 'grab'). Again, there is an overlap with posture verbs which might be used in basic location and existential constructions. Posture to location/existence being a known grammaticalization path (e.g. Newman 2002), we favored Posture over Location if posture was listed in the translation. The Location category occurs in 35 languages and 50 tokens, and stands thus far behind Motion and Posture. It is notable that only six languages[23] have neither Motion, Posture nor Location.

The Frequent Verbs category includes a heterogeneous group of frequently attested verbs already noticed in the literature. They have highly specific semantics – 'die', 'kill', 'sleep', 'cry', and 'eat' – and their frequencies are given in Table 10. The first two ('die' and 'kill') are almost as frequent as the entire category of Location, and each appear in 30 languages.

The All Others category includes all verbs which did not fit into one of the four previous categories and contains a variety of semantics. We looked through the semantics of individual verbs and created five *post hoc* subcategories that could accomodate approximately half of the tokens: Destruction, Property, Transfer, Emission, and Perception verbs (Table 11).

Destruction verbs (e.g., 'hit', 'pierce', 'break') are the most frequent, with 42 tokens across 25 languages. Note that 'kill' and 'die', which are grouped with Frequent Verbs, could be included in this category.[24] Property verbs (e.g. 'be big', 'be dirty') are

Table 10: Frequent Verbs of suppletive alternations by languages.

Frequent Verbs	Languages (*N* = 103)
'die'	30
'kill'	30
'sleep'	15
'cry'	13
'eat'	10

23 Caddo (Caddoan), which only alternates the 'be' verb, Akabea (Great Andamanese), Magdalena Peñasco Mixtec (Otomanguean), Huastec (Mayan), Bauni (Sko), and Krongo (Kadugli-Krongo).

24 François (2019: 352) also notes this category in Hiw and Lo-Toga, which he terms "intense physical impact" verbs, and in which he includes 'die' and 'kill'.

Table 11: Meaningful semantic subcategories in suppletive alternations in "All Others".

Subcategories	Languages (*N* = 103)	Tokens (*N* = 131)
Destruction	25	42
Property	15	39
Transfer	15	19
Emission	15	23
Perception	5	8

also common, with 39 tokens across 15 languages. Another 17 languages could be added to this count, if suppleting adjectives are considered in addition to verbs. The Property category has been noted by Booker (1982), although she only comments on 'big' and 'small'. Transfer verbs (e.g. 'give', 'get') have only 19 tokens across 15 languages, but several Motion verbs like 'throw' or 'bring' could be included in this category.

Verbs of Emission (e.g. 'talk', 'speak') and Perception (e.g. 'see', 'hear') respectively amount to 23 tokens across 15 languages and eight tokens across five languages. Although basic and common verbs in most languages, they hardly supplete for number when compared with verbs from other categories, even Destruction and Property. This discrepancy points to at least a partially semantic-based, rather than a purely frequency-based, explanation for verbs targeted by Sg-Pl alternation, to which we return in the discussion. We were unable to define further categories with more than two lexical items and occurring in more than a handful of languages.

5 Local results

As shown in the global results in Section 4, Sg-Pl alternation is present in a fourth to a third of our sample, but with pronounced distributional differences among macroareas. We now turn to local clusters of the phenomenon and examine the roles of inheritance (5.1) and areality (5.2).

5.1 Inheritance

The extreme geographic skew toward North America raises the prospect that this feature may be especially present in certain language families. The language families in North America could have inherited this feature, either from the known proto-

languages, or perhaps from a theoretical common ancestor of North American languages that lies beyond the limit of the comparative method.[25]

At the continent-wide level, the inheritance model is a poor explanation of the difference between North America on the one hand and Central and South America on the other. If Sg-Pl alternation is a long-term stable feature reflecting deep shared ancestry, then the sharp typological shift that occurs south of the Chihuahuan Desert is a problem. If North American families inherited this feature from the distant past, then that process should affect languages further south as well.

The inheritance model also cannot explain the distribution within known families in North America. As noted in Section 4.1, Sg-Pl alternation is often unevenly distributed within a family. Though this feature has not been a focus of reconstruction efforts, we will briefly review two detailed studies on Uto-Aztecan (Haugen and Everdell 2015) and Salishan (Kinkade 1981), and a less detailed study on Wakashan (Inman 2021). These, as well as the results of our survey, suggest that most alternating verbs have developed after these languages differentiated.

The Uto-Aztecan family provides the strongest potential example for widespread inheritance of Sg-Pl alternation. Of the eight Uto-Aztecan languages in our sample, six have Sg-Pl alternation, including all languages from the Northern branch (Cahuilla, Hopi, Comanche, and Northern Paiute), and two from the Southern branch, members of the Cahitan (Yaqui) and Piman (Tohono O'odham) subbranches. Only the Aztecan languages in our sample lack the phenomenon (Pipil and Tetelcingo Nahuatl). A look in grammars of Southern, non-Aztecan languages outside of our sample confirms the presence of Sg-Pl alternation in Southeastern Tepehua [sout2976] (Tepehuan subbranch; Willett 1991: 44, 125, 171), Huarijio [huar1255] (Tarahumara-Guarijo; Félix Armendáriz 2005: 116), and Huichol [huic1243] (Corachol; Bierge 2017: 130).

Despite this abundance of Sg-Pl alternation within the family, Haugen and Everdell (2015) reconstruct only two verb pairs in the proto-language: 'die' and 'kill'. While the phenomenon was thus present to at least some degree in Proto-Uto-Aztecan, it has greatly developed in many languages of the Northern branch (Comanche, Hopi, and Northern Paiute all have at least 18 suppleting verb pairs) as well as in (Southern but geographically northern) Yaqui (20 verb pairs), while it remained small or dwindled out elsewhere. Interestingly, the branch where Sg-Pl alternation is not attested, Aztecan, is entirely spoken within Central America.

25 According to modern genetics and anthropology, North and South America are the most recently populated continents. They could consequently share linguistic features due to a more recent common ancestry than other parts of the world. The mismatch between this hypothesis and the reality of extreme linguistic diversity in the Americas has long been a puzzle for linguists, but we do not rule out *a priori* that some typological similarities may result from this comparatively recent settlement.

The distribution of the phenomenon in Salishan is very similar. Kinkade (1981) documents between six and 23 verbal pairs in seven Salishan languages, all in the Interior or Tsamosan branches. He did not find many examples among Central Salish languages or in Bella Coola, a distribution also reflected in our sample. Kinkade finds only a few cognates among the Interior languages, and in only a few cases was he able to trace the alternation to the phonological reduction of an attested regular plural process. He concludes that this phenomenon, despite its abundance in some languages, is relatively recent within the family, and proposes that it was borrowed, possibly from Athabaskan.

Though less robustly documented, the same pattern of recent spread seems to hold true for Wakashan. Inman (2021) reports on plural processes in Nuu-chah-nulth verbs and nouns and briefly compares them to other Wakashan languages. Nuu-chah-nulth has four suppletive pairs,[26] and Kwak'wala is reported to have two. Alternating pairs are (so far) not documented for the rest of the family. Similar to the other families discussed above, there are no cognate suppletive roots and the feature seems to be a more recent development, rather than inherited from Proto-Wakashan.

Finally, as alluded to in Section 4.1, four out of the five Otomanguean languages in our sample have Sg-Pl alternation, which represents a disproportionate number of the Central American languages with the phenomenon. However, only the semantics 'sit' is present across three of them (Nellis and Nellis 1983: 322; Suárez 1983: 161; Veerman-Leichsenring 1991: 119), and there is no obvious cognacy among the forms (Eric W. Campbell, pers. comm., March 2023).

Sg-Pl alternation can undoubtedly be inherited. However, even in cases in which it has persisted in a family through time, the individual verb pairs are not necessarily preserved, and cognates within families are very rare.

5.2 Areal spread

Inheritance being insufficient to describe the distribution of Sg-Pl alternation, we now turn to borrowing and linguistic areas as an explanation, as proposed by Kinkade (1981) for Salishan and by Mithun (2022) for several unrelated Californian languages.

The mechanism proposed by Mithun (2022) is that second language speakers, accustomed to Sg-Pl alternation in their L1, overuse hypernyms in their L2 to signify plurality, which over time leads to the spread of Sg-Pl alternation within a given

26 Note that we have not counted 'swim', due to the uncertainty of whether or not it is in a hypernym/hyponym relationship akin to English *swim* and *school*.

linguistic area. The same process, of course, happens in the other direction, with L1 speakers (of a language without Sg-Pl alternation) learning to speak an L2 (with Sg-Pl alternation), and bringing this feature home to their L1. Importantly, despite this multilingual mechanism, the data reveal no form borrowing across languages in any of the areas investigated below.

We will examine this process of areal spread by looking at the specific semantics of suppletive verb pairs in five different regions: the well-established linguistic areas of the Pacific Northwest, California, and the Pueblo area, all of which are in North America, as well as the dense clusters of Sg-Pl alternation revealed in the Guaporé-Mamoré region and Eastern Papua. Recall that we only collected *underived* verb pairs, which might diminish the chance of having similar semantics across languages. One can easily imagine that a language derives an additional suppletive pair with regular valency changing markers to copy alternation occurring in two underived suppletive pairs in the neighboring language.

We are also only presenting suppletive alternations within each area. As we have noted, suppletion is the most typical form of alternation worldwide, and while this restriction makes our data easier to interpret, it runs the risk of missing contact events that occur between suppletive pairs and regular patterns.[27] Comparing suppletive pairs will however capture any Sg-Pl alternation that spreads according to an areal mechanism of hypernym generalization, as outlined in Mithun (2022).

5.2.1 Pacific Northwest

The Pacific Northwest has long been recognized as a linguistic area (e.g. Boas 1898, 1911c; Jacobs 1937; Thompson and Kinkade 1990), but its boundaries are not well-defined. We will consider the set of languages provided in Beck (2000). Like much of North America, these languages frequently have root suppletion. Interestingly, the semantics of alternating verbs is frequently shared between neighboring languages, but not necessarily across the entire region. Figure 5 is a map of all the Pacific Northwest languages in our sample. Table 12 presents the suppleting verbal pairs. Semantics that are shared between neighbors or near-neighbors, regardless of their family, are bolded.

The semantics of most verbs pairs shared between neighbors are globally typical of Sg-Pl alternation ('go', 'run', 'sit', 'stand', etc.); however, some are regionally specific, such as 'jump' in Alsea-Yaquina and Hanis, and 'throw' in Klamath-Modoc and Karok. Some sub-regions share more verb semantics than others: The alternating pairs in the northern end of the linguistic area (roughly Eyak to Southern-Coastal

27 An example is the verb pairs for 'dance' in Alsea-Yaquina (where the pair is suppletive) and in neighboring Hanis (where the pair is formed by a regular pattern).

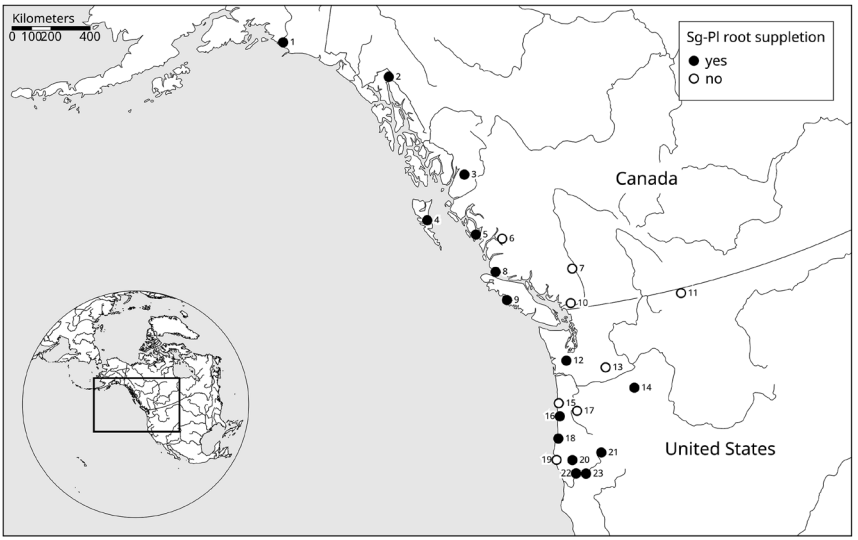

Figure 5: Languages in the Pacific Northwest. Those without Sg-Pl alternation are: 6 Bella Coola (Salishan), 7 Lillooet (Salishan), 10 Halkomelem (Salishan), 11 Kutenai (isolate), 13 Northwest Sahaptin (Sahaptian), 15 Central Kalapuya (Kalapuyan), 17 Molale (isolate), 19 Siuslaw (isolate).

Table 12: Suppletive alternations in the Pacific Northwest.

#	Language	Family	Suppletive verb pairs
1	Eyak	Na-Dene	MOTION: **go**, **run** POSTURE: **sit**, **stand**, **lie (prone)**, extend
2	Tlingit	Na-Dene	(CAUSED) MOTION: **walk**, **run**, **fly** POSTURE: **be sitting**, **sit down**, **stand**, **lie** FREQUENT: **cry**
3	Nisga'a	Tsimshianic	(CAUSED) MOTION: **go/walk**, go out, go up a slope, **leave**, be/come back, **run**, flee, ride, **put away**, **take**, throw in a fire POSTURE: **sit**, **stand**, **lie**, hang CAUSED LOCATION: hold/carry FREQUENT: **die**, **kill**, **cry**, **eat** OTHER: **grow**, be small
4	Southern Haida	Haida	MOTION: **go**, **fly** POSTURE: **sit** FREQUENT: **kill**
5	Southern-Coastal Tsimshian	Tsimshianic	(CAUSED) MOTION: **go**, **go out**, **leave**, **run**, start off, escape, drift away, **put away**, put down, **take in** (CAUSED) POSTURE: **sit**, **stand**, **lie**, place upright FREQUENT: **die/kill**, **eat** OTHER: be dirty, be easy, be ripe (of a rain cloud), be sour, be strong, tear out and turn over, **grow**, name, wake up, wipe

Table 12: (continued)

#	Language	Family	Suppletive verb pairs
8	Kwak'wala	Wakashan	Posture: **sit**, **stand**
9	Nuu-chah-nulth	Wakashan	Motion: **run**, **fly** Frequent: **cry**, sleep
12	Upper Chehalis	Salishan	(Caused) Motion: run, walk, fall, drop, throw Posture: sit, **stand**, lie/put down Caused Posture: grab Frequent: **kill**, **cry**
14	Clatsop-Shoalwater Chinook	Chinookan	Posture: **stand** Location: be (there) Frequent: **die**, **kill**, **cry**
16	Alsea-Yaquina	isolate	Motion: dance, enter, enter water, **jump** Posture: **sit**, **stand** Frequent: **die**, **cry**, **sleep** (verb 1), **sleep** (verb 2) Other: agree
18	Hanis	Coosan	Motion: **fly/jump**, travel Posture: **sit**, **lie** Frequent: **die**, **kill**, **sleep** Other: do, speak
20	Takelma	isolate	Posture: **sit**, **stand**
21	Klamath-Modoc	Isolate	(Caused) Motion: **run**, **throw** Posture: **sit/live/stay**, **stand**, **lie down**, float Frequent: **die**, **kill** Other: give, hit/shoot
22	Karok	Isolate	(Caused) Motion: **swim**, fly, climb, **throw** Posture: **sit**
23	Shasta	Shastan	Motion: **run** (verb 1), **run** (verb 2), walk, **swim**, jump, fall, go out to camp, move suddenly, move along Posture: **sit**, **stand**, lean Caused Location: hold Frequent: sleep Other: look, bite, hunt/dig/mine

Tsimshian note the relative phylogenetic homogeneity) repeat quite frequently language-to-language, and pairs are shared to a slightly lesser degree in the south (roughly Clatsop to Shasta, note the relative phylogenetic diversity). In the central region of the Pacific Northwest, traditionally thought of as the core of the linguistic area (the Salishan, Wakashan, and Chimakuan families, or roughly Bella Coola to Upper Chehalis, see Beck 2000), there are more languages lacking Sg-Pl alternation altogether.

In addition to these similarities between neighboring languages, there is also significant diversity in the region. Despite belonging to the same family, Nisga'a and Southern-Coastal Tsimshian have a large number of non-shared suppletive stems. In the southern sub-region, there are a large number of idiosyncratic suppletive alternations. Overall, the picture suggests both areal spread and independent innovation.

5.2.2 California

Much like the Pacific Northwest, the exact boundaries of the California linguistic area are somewhat fuzzy, and many sub-areas have been proposed (Northern California, Northwest California, Clear Lake, the South Coast Range, and Southern California-Western Arizona, see Hill 2019). Since our sample is relatively sparse, we will report on Sg-Pl alternation among all languages spoken in what is today California, plus Northern Paiute (due to its proximity to Washo). Note that Karok and Shasta are also present in the Pacific Northwest grouping given above. Our data are presented in Figure 6 and Table 13.

As in the Pacific Northwest, there are regions in California where languages share verbal semantics, although these are often the typical semantics of the phenomenon as a

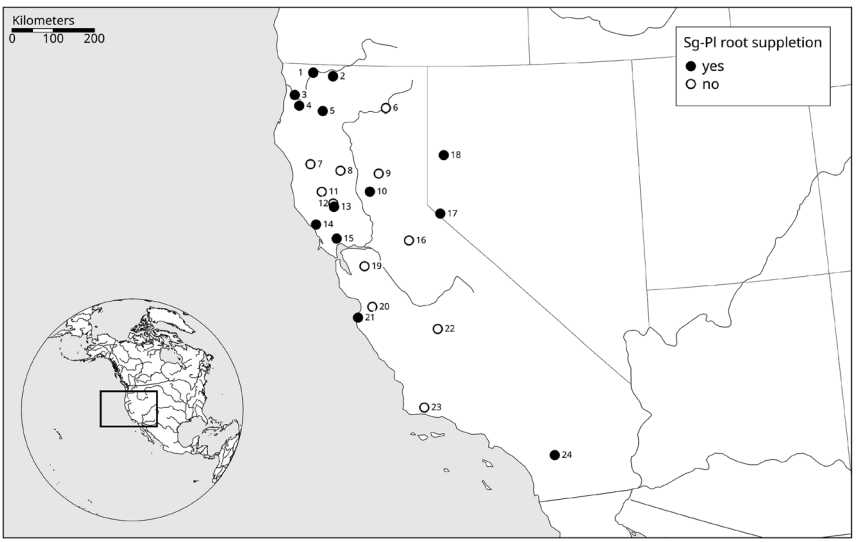

Figure 6: Languages in California. Those without Sg-Pl alternation are: 6 Achumawi (Palaihnihan), 7 Northern Yukian (Yuki-Wappo), 8 Wintu (Wintuan), 9 Northwest Maidu (Maiduan), 11 Wiyot (Algic), 12 Lake Miwok (Miwok-Costanoan), 16 Southern Sierra Miwok (Miwok-Costanoan), 19 Salinan (isolate), 20 Mutsun (Miwok-Costanoan), 22 Northern Yokuts (Yokutsan), 23 Ineseño (Chumashan). Note: Northwest Maidu has however three bound (verbal) stems (*wO-*, *bO-*, and *he-*) which regularly alternate by vowel lengthening and which we considered in feature SgPl.19 in the supplementary materials (see the Data availability statement) on Sg-Pl alternating morphology. Their semantics is unclear.

Table 13: Suppletive alternations in California.

#	Language	Family	Suppletive verb pairs
1	Karok	Isolate	(Caused) Motion: **swim**, fly, climb, throw Posture: **sit**
2	Shasta	Shastan	Motion: **swim**, **run** (verb 1), **run** (verb 2), walk, jump, fall, go out to camp, move suddenly, move along Posture: **sit**, **stand**, lean Caused Location: hold Frequent: sleep Other: look, bite, hunt/dig/mine
3	Yurok	Algic	Motion: arrive Posture: **sit**, **stand** (inanimate)
4	Hupa-Chilula	Na-Dene	(Caused) Motion: **run**, move, move something somewhere, **carry** Posture: **stand**, lie down, lie motionless, have a position Location: stay/live Frequent: kill
5	Chimariko	isolate	(Caused) Motion: **run**, **carry** (a firm object), **carry** (a load), **carry** (a long object), **carry** (a small object) Posture: **sit**
10	Yana	isolate	Motion: go (M), go (F), **go/move**, **run** Posture: **sit**, sit/stay/live Other: laugh, awaken, be strong, be long
13	Southeastern Pomo	Pomoan	(Caused) Motion: **run/fly/flow**, **walk/go**, **swim**, roll/fall, float, **crawl**, come to rest on water, come to rest (long object), plop down, place X on ground, place X on surface Posture: **sit** (non-long object), sit down, **stand** (long object), **lie** (long object)
14	Kashaya	Pomoan	(Caused) Motion: **go**, **crawl/swim**, **carry** (any-shaped object), **carry/hold** (long object), **carry/hold** (non-long object), implant (long object), put (long object), put (non-long object) Posture: **sit** (on ground), **sit** (on platform), **stand** (on feet), **stand** (as tree), **lie**, hang Frequent: **kill** Other: shoot, give
15	Wappo	Yuki-Wappo	(Caused) Motion: **go away**, come out, **run**, **walk**, **carry** Posture: **sit**, **stand**, **lie** Frequent: die, **kill**, sleep Other: sing
17	Washo	isolate	Motion: run, **fly**, swim, fall off, fall over Posture: **sit**, **stand**, **lie**
18	Northern Paiute	Uto-Aztecan	(Caused) Motion: go, come, enter, return, travel, **fly**, put Posture: **sit**, **stand**, **lie** Frequent: die, kill Other: talk, give, make a hole in, open
21	Esselen	isolate	Caused Motion: **bring** (cause to be **carried**)

Table 13: (continued)

#	Language	Family	Suppletive verb pairs
24	Cahuilla	Uto-Aztecan	(Caused) Motion: fall, put, take Posture: lying flat Location: exist, stay Frequent: die Other: shoot, light a fire

whole. However, there are also more idiosyncratic verbs (e.g. 'carry' in Hupa-Chilula and Chimariko, as well as in Kashaya and Wappo, and 'crawl' in Southeastern Pomo and Kashaya). Once again the picture is mixed, revealing both cross-linguistically shared semantics and considerable variation.

5.2.3 Pueblo

The Pueblo linguistic area belongs to a well-established cultural area of the historic pueblos of the American Southwest (Eggans 1979). It consists of Zuni (isolate), Hopi (Uto-Aztecan), Navajo (Athabaskan-Eyak-Tlingit), and the Kiowa-Tanoan and Keresan families (Campbell 1997: 339), although it should be noted the Navajos arrived in the area only just before the Spaniards (Eggans 1979). From the perspective of Sg-Pl alternation, these languages do not look markedly different from neighboring languages. However, we focus here on only the Pueblo languages, since they are known to have been in intensive contact. These languages also have some of the largest inventories of verb pairs.[28] Because of the tight-knit nature of this linguistic area, we consider every language to be near-neighbors (Figure 7 and Table 14).

The languages with relatively small inventories (Arizona Tewa and Navajo) share all or almost all of their verb pairs with their neighbors, including both more common semantics like 'sit' and 'stand' and relatively less common semantics like 'throw'. Languages with very large inventories (Hopi and Western Keres) share some semantics with their neighbors but have a much larger inventory of idiosyncratic pairs.

5.2.4 The Guaporé-Mamoré region

The cluster of languages around the Brazilian state of Rondônia is a part of the proposed Guaporé-Mamoré linguistic area (Crevels and Voort 2008), and stands out

28 Western Keres has the largest inventory of any language in our sample. Hopi has the third-largest inventory, and relatively nearby Comanche (still outside of the Pueblo linguistic area) has the second-largest inventory.

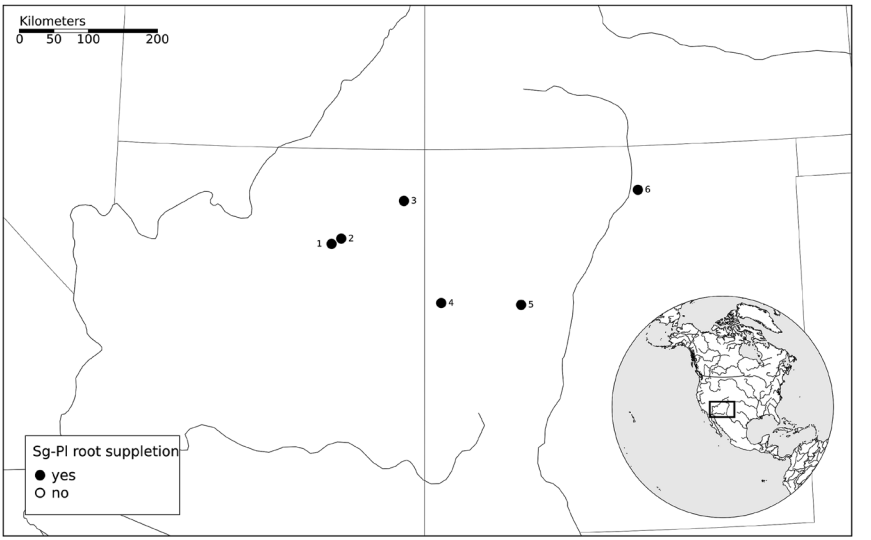

Figure 7: Languages in the Pueblo linguistic area.

Table 14: Suppletive alternations in the Pueblo linguistic area.

#	Language	Family	Suppletive verb pairs
1	Hopi	Uto-Aztecan	(Caused) Motion: **arrive**, **enter**, come out, **fall/drop**, **run**, **climb up**, **climb down**, dance, wander, escape, **throw**, **put**, put into a rigid/ enclosed container, put on top, carry (animate), carry (inanimate), bring back (animate), fetch, get extracted Posture:: **sit**, **stand** Frequent: **die**, **kill**, **cry**, **sleep**, **eat** Other: **receive**, **give**, break (linear segments), break/crack, dress, talk about, harvest/pick from tree, laugh
2	Arizona Tewa	Kiowa-Tanoan	(Caused) Motion: **fall**, **throw away** Posture: **sit**, **lie**, **stand up** Other: **get**, **hunt/dig/mine**
3	Navajo	Na-Dene	(Caused) Motion: **go**, **run**, **throw**, lead Posture: **sit**, **lie down** Other: **die**, **kill**
4	Zuni	isolate	(Caused) Motion: **run away**, **take**, **put against**, **put down** Posture: **sit**, **stand**, **lie** (verb 1), **lie** (verb 2), hang (Caused) Location: be among, be inside, hold Frequent: **kill**, **sleep** Other: **give**
5	Western Keres	Keresan	(Caused) Motion: **go**, **arrive**, **enter**, leave, walk, **fall in**, **fall off**, **climb**, pass by, step down, fly, rise, **take**, **throw**, **throw away** Posture: **lie down**, **lean**, **stand up**, **be standing**, float

Table 14: (continued)

#	Language	Family	Suppletive verb pairs
			LOCATION: dwell FREQUENT: **die**, **cry**, **sleep**, spend the night, **eat** OTHER: say, speak up, hear, listen, look for, buy, lose, find, be closed (eyes), finish, have it closed, have luck, know, know how, tend, nurse, release, wait for, wake up, win, grow, grow old, be lost, be enclosed, be stuck, be big, be fat, be hungry, be strong, be a fast runner, be talkative, be pregnant, be born, be first, drink
6	Taos Northern Tiwa	Kiowa-Tanoan	*Unknown (mentioned in grammar but no list given)*

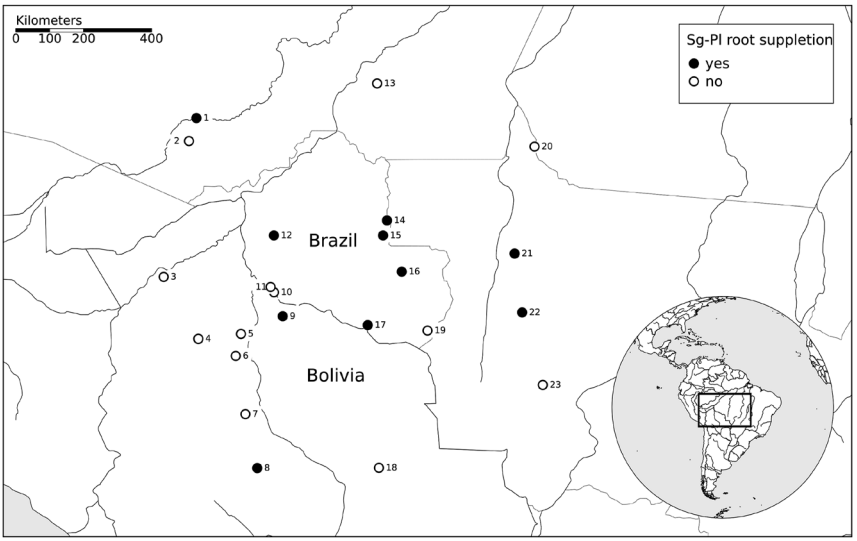

Figure 8: Languages in the Guaporé-Mamoré region. Those without Sg-Pl alternation are: 2 Apurinã (Arawakan), 3 Ese Ejja (Pano-Tacanan), 4 Cavineña (Pano-Tacanan), 5 Cayubaba (isolate), 6 Movima (isolate), 7 Trinitario-Javeriano-Loretano (Arawakan), 10 Kanoê (isolate), 11 Djeoromitxí (Nuclear-Macro-Je), 13 Pirahã (isolate), 18 Lomeriano-Ignaciano Chiquitano (Chiquitano), 19 Lakondê (Nambiquaran), 20 Mundurukú (Tupian), 23 Parecís (Arawakan).

in our sample as a hotspot of Sg-Pl alternation in South America. This is one of the clearest clusters of the phenomenon among a phylogenetically diverse group outside of North America (Figure 8 and Table 15).

Table 15: Suppletive alternations in the Guaporé-Mamoré region.

#	Language	Family	Suppletive verb pairs
1	Madi	Arawan	(Caused) Motion: **fall**, put in water, put inside, put on ground/surface, take out, **throw** Posture: sit/stay, **stand**, **lie on ground**, **lie on surface**, lie on water, hang, tie onto a hook (Caused) Location: be inside, hold in hand Frequent: kill Other: roast/fry, pierce, **be big**
8	Yuracaré	isolate	(Caused) Motion: **go**, go in, **fall**, **take**, **throw** Posture: **lie (down)** Frequent: die
9	Itonama	isolate	Motion: **fall**
12	Wari'	Chapacuran	(Caused) Motion: **go**, **arrive**, **fall**, fly, **run**, **take**, **throw** Posture: **stand**, **lie down** Location: be at/live Frequent: **die**, **cry** Other: **be big**, be stuck
14	Karo	Tupian	Motion: **run**, come back Frequent: **cry**, sleep
15	Gavião Do Jirapaná	Tupian	Motion: **go** Frequent: **die**, **kill** Other: beat, knock down, make
16	Kwaza	isolate	(Caused) Motion: **move**, cross Posture: **sit**, **lie down** Other: end, sting on head
17	Mekens	Tupian	Motion: **go**, **go/come**, walk Posture: **sit**, **stand**, **lie**
21	Rikbaktsa	Nuclear-Macro-Je	Motion: **run** Frequent: **kill**
22	Irántxe-Münkü	isolate	Location: live

Though the inventory of verb pairs is much smaller in this region than it is in the Pueblo linguistic area, a very similar pattern emerges: Small inventories tend to be fully or nearly-fully shared with neighboring languages (e.g. Rikbaktsa, Mekens, Itonama), while larger inventories have more idiosyncratic semantics (e.g. Wari', Kwaza).

5.2.5 Coastal Eastern Papua

Because of the skew of our sample, our data in the Americas are much more densely sampled than in the other four macroareas. However, Papunesia is a small macroarea with high linguistic diversity and thus provides us a relatively dense sample of the island of Papua and nearby New Britain. It is also the macroarea with the second

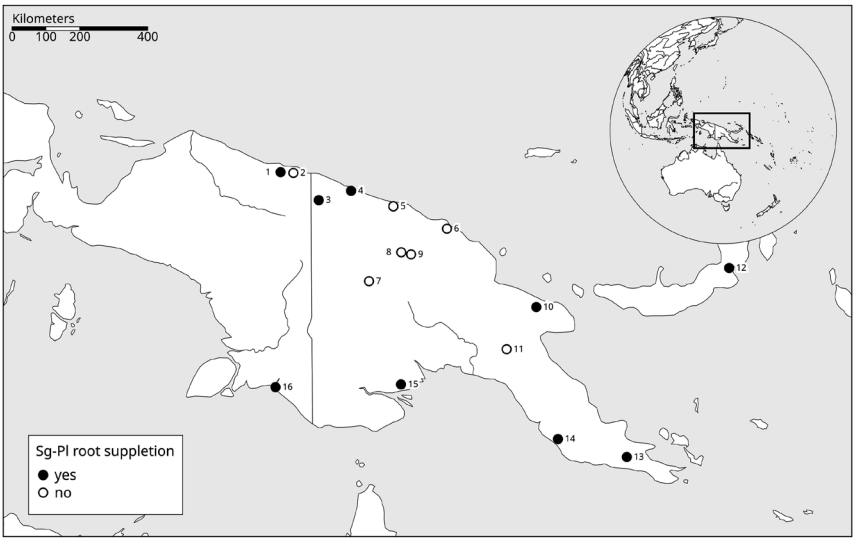

Figure 9: Languages in Eastern Papua and New Britain. Those without Sg-Pl alternation are: 2 Sentani (Sentanic), 5 Bukiyip (Nuclear Torricelli), 6 Taiap (isolate), 7 Duna (isolate), 8 Alamblak (Sepik), 9 Yimas (Lower Sepik-Ramu), 11 Menya (Angan).

Table 16: Suppletive alternations in Eastern Papua and New Britain.

#	Language	Family	Suppletive verb pairs
1	Nimboran	Nimboranic	(Caused) Motion: climb, **put**, take, throw
3	Imonda	Border	(Caused) Motion: **follow**, **put**, **put up**, put into fire, stick into (Caused) Posture: sit, sit down, lie, stand, hang up (Caused) Location: hold, hold/possess Frequent: sleep, eat Other: **get**, give, cut, hear, see, fill in, make netbag, dislike
4	Bauni	Sko	Frequent: die Other: **get**
10	Yau-Nungon	Nuclear Trans New Guinea	Caused Motion: **follow**, take/pick up, take away Other: **beat**
12	Sulka	isolate	(Caused) Motion: arrive, **put (away)** Other: **get**, cook, dig, scrape
13	Daga	Dagan	Caused Motion: throw, **put down** Frequent: **kill/hit**, Other: **get**
14	Grass Koiari	Koiaran	Caused Motion: **put** Posture: **sit** Other: **get**, **become/be**, defecate, urinate
15	Tabo	isolate	(Caused) Motion: appear, **ascend**, descend, jump, **put**, take Posture: **sit** Location: stay

Table 16: (continued)

#	Language	Family	Suppletive verb pairs
16	Marind	Anim	FREQUENT: **kill/strike** OTHER: grow, scream, small (CAUSED) MOTION: be in movement, come, go/leave, **climb**, run, run away, **put** (horizontal item), **put** (vertical item), put inside, throw POSTURE: **be sitting**, **sit down**, be standing, stand up, be lying, lie down CAUSED LOCATION: catch, grasp (horizontal item), grasp (vertical item) FREQUENT: die, cry, eat OTHER: **hit**, shoot, call (someone's name), **become**, father (a child), give birth to, plant coconut

most frequent occurrence of Sg-Pl alternation. We can thus report on the phenomenon in Eastern Papua, where the same pattern of shared semantics and internal diversification recurs (Figure 9 and Table 16).

These languages mostly share the semantics of 'put', and the less common semantics of 'get' (the only other similar verbs are 'get' in Arizona Tewa and 'receive' in Hopi, both in the Pueblo region, see Section 5.2.3). The only two occurrences of suppletion for the word 'follow' in our sample occur in this region, in Yau-Nungon and Imonda, and though these two languages are relatively far apart, there may be intervening speech communities not in our sample which would reveal a chain of languages with alternating verbs for 'follow'. Also, unlike other regions, the above languages lack common spontaneous motion verbs such as 'go' and 'come' and have very few posture verbs, except for Marind and Imonda. Recall also that transitive-only languages have almost only been observed in this region, which also has no intransitive-only languages (see Section 4.2). A more targeted study in Papunesia would likely be fruitful in uncovering more details about contact-induced Sg-Pl alternation.[29]

There may be other clusters of languages displaying Sg-Pl alternation in regions where our sampling is too sparse to make strong observations. Possible locations include Northwestern Australia, the East African Rift Valley, and the Caucasus.

Regardless of the region, suppletive alternations overwhelmingly fall into the semantic domains listed in Section 4.4. The areal picture that emerges from the individual regions is a mixture of two trends: shared verbal semantics, likely through

29 We do not suggest that Papunesia is homogenous. For instance, the two Vanuatu languages studied in François (2019: 349) do alternate 'go' (in Hiw) and several posture verbs (in Hiw and Lo-Toga), and not 'get'. In addition, Antoinette Schapper (pers. comm., May 2023) reports that in the west some languages alternate motion verbs like 'go' and 'come' and some display Sg-Pl alternation in intransitive verbs only.

direct contact, and significant internal variation, especially once the inventory of alternations becomes large (more than eight or so verb pairs). The first trend indicates direct calquing of the semantic organization of the lexicon, as noted by Mithun (2022) for California. This frequently happens at the sub-areal level, with close neighbors sharing semantics which often do not extend to the area as a whole. The second trend steps in once the feature is sufficiently present in a language: Sg-Pl alternation can then diffuse across the lexicon without the need for semantic calquing. We finally want to reiterate the absence of form borrowing across any of the areas investigated.

6 Discussion

6.1 Syntax and semantics together

The strong global preference for Sg-Pl alternation to follow an S/P pattern, and to target the semantics of Motion, Posture, and Location, together suggest a cognitive or a functional motivation at least partially driving the phenomenon. Durie (1986: 357) and Mithun (1988: 214) propose that the S/P pattern arises from the semantics of the targeted entity, namely the "affected argument." However, many verbs with highly affected patients such as 'hit', 'kick', and 'slap' are infrequent in our sample. The presence of a highly affected argument is thus insufficient to explain the semantic profile of alternating verb pairs. We instead believe that a significant motivation is the visual configuration of central entities in Motion Events.[30]

Motion, Posture, and Location events are all Motion Events in the sense of Talmy (1972), in that they all involve a Figure being located or moving with respect to a Ground. The Figure is thus very central, and has an asymmetric relationship with the Ground: the Figure is typically more movable and smaller than the Ground, of greater concern, and less immediately perceivable but more salient once perceived (Talmy 2000).

By definition, Motion Events are spatially oriented and thus highly visual. If the Figure contains more than one entity, their configuration (grouped, dispersed, etc.) is consequently more salient than that of the arguments of basic perception and emission events. The significance of configuration in Motion Events is so great that Motion lexemes frequently target specific configurations even in languages with no Sg-Pl alternation. For instance, the English word *scatter* specifies that the moving P is

30 During the final revisions of this manuscript, we became aware that François (2019) had already proposed a similar hypothesis. He (p. 352, 359) discusses spatial configuration, namely the salience of an "individual versus collective configuration" for certain verb semantics, and especially for posture.

both plural and has a resulting dispersed spatial configuration, which distinguishes it from *drop*, which lacks both numerical and configurational specification of its arguments. (Note that *drop* and *scatter* are not an example of a Sg-Pl verb pair, as only *scatter* specifies number. Further, transitive *scatter* must be intentional, while *drop* may be unintentional.) Number, as a necessary component of configuration, is thus also more likely to cause lexical differentiation among Motion Events. What sets apart Sg-Pl alternation is that the lexemes are pairs contrasting number only.

In transitive Motion Events, it is typically the P argument (and not the A) which is the Figure, and according to which the alternation is triggered. The A argument is however sometimes the Figure as in '**A** crosses a bridge' and '**A** lives at X'. Interestingly, transitive 'cross' and 'live at' are two of these A-triggered verb pairs, along with 'chase', where A is also moving (following the Figure).[31]

Spatial configuration may also be relevant for three of the "Frequent Verbs". Both 'die' and 'kill' typically involve a significant postural change in the entity passing away. (This comes of course in addition to the cognitive salience of the number of entities dying.) The verb 'sleep' is typically associated with the lying posture (CLICS lists colexifications of 'lie down' or 'lie (rest)' with 'sleep' in 225 languages; Rzymski et al. 2020). By contrast, other semantic classes like cognitive and mental states (e.g. 'think', 'dream') and basic perception or sensation (e.g. 'see', 'hear', 'feel') lack salient spatial configuration, which may explain their infrequency in our list of verb pairs.

The spatial configuration of objects is probably not the only motivation. The cultural significance of certain activities may be important, as already noted by Mithun (1988: 232). This is true for both 'sleep' and 'eat', which can have culturally distinct meanings for an individual as opposed to a collective (e.g. a family or village). Other somewhat-common categories like Destruction and Property verbs do not seem to be straightforwardly related to either culture or configuration.

Note that Mithun (2022: 375) considers not only number, but also animacy, shape, and/or consistency among the features relevant to the verb semantics in the Californian area. Our data suggest that animacy, shape and consistency of the entity are not particularly relevant in the verb semantics globally, since they were only present in Yana, Chimariko and Kashaya (North America), Marind (Papunesia), and Kaingang (South America).[32] We therefore believe that these three parameters are areally relevant lexico-semantic classifications in the Californian languages, and these

31 The other A-triggered verb pairs in our sample are 'hit with a stick', 'see', 'need', 'strengthen', 'dislike', 'make netbag', 'teach', 'talk about', and 'eat'.

32 Burushaski (Eurasia) and Choctaw (North America) could arguably also belong to this set. Both have a suppletive pair for 'collect/pick up fruit', for which the shape of the P argument may be relevant.

semantics thus are sometimes also present among Sg-Pl verb pairs. In other words, we think these are two separate typological phenomena, independently present in California. Our sample is however not as dense as hers in the region, and our methodology excludes derived and compositional verb pairs. Thus, where Mithun (2022: 387–8) counts several verbs (still pointing to some compositionality), we count only one alternating root.

Finally, we cannot rule out the possibility that our definition, which specifically targets argument rather than event number, affected our semantic results. The most relevant number for events like 'hit' and 'beat' may be the repetition of the event, while for Motion Events, it is the argument number. The semantics of lexical alternation strictly according to event number remains to be systematically surveyed.

6.2 Explaining the macroareal distribution

The link between the local influence driving the adoption of particular alternating pairs (Section 5.2) and the macroareal level (Section 4.1), at which North America stands out as a whole, remains somewhat unclear. Previous researchers have proposed various mechanisms for the overwhelming presence of Sg-Pl alternation in North America, but we find all of these proposals problematic.

Mithun (1988: 228) suggests that this feature is linked to number marking in North America, which is typically verbal rather than nominal, and expresses distribution rather than simple plurality. Under certain circumstances morphological distributivity can shift to plurality when it becomes lexically incorporated into the verb, a process she documents for some Iroquoian languages. From our survey, this pathway appears to be fairly atypical. As explained in Section 4.1, suppletion is far more frequent than lexically limited regular patterns: 75/109 languages have only suppletive pairs, only 5/109 languages have only regular patterns, and regular patterns overall are less common than suppletion. If (lexicalized) distributive morphology generates lexical alternations, it appears to be a relatively uncommon process. However, it remains possible that a category of distributive number may contribute to a cognitive organization of number as a verbal (rather than a nominal) property, which then would contribute to the formation of Sg-Pl alternation.

This is related to another proposal, namely that the absence of nominal number marking promotes number marking (including suppletion) on the verb. One such proposal is given by Frajzyngier (1985: 99–100) (criticized in Durie 1986: 367), who suggests an inverse correlation between productive nominal number marking and the marking of a plural P role on the verb. To our knowledge, no hypotheses of this sort have been tested on typological data at scale. To do so, "lack of nominal number"

would need to be narrowly defined: for example, is obligatory number marking for humans and optional number marking for animals and inanimates considered to be a lack of nominal number or not? For Frajzyngier, the association is with productive nominal plurals, and not the presence of any plural marking at all. Our survey did not keep track of the presence of distributive number nor of obligatory nominal number marking. We are inclined to agree with Durie (1986) and consider these phenomena to be of different natures, but this remains to be systematically investigated.

Another explanation that is sometimes proposed for various typological phenomena is the mode of subsistence (agriculture vs. hunter-gatherers) and the size of the speaker population. Given the presence of Sg-Pl alternation in North America and the historic dominant subsistence mode on that continent (which in many places depended on following a "yearly round" of seasonally-available resources, rather than settled agricultural plots, see Smith 2006), it is worth addressing briefly. We are fairly confident that subsistence mode and population size do not play a strong role in driving this distribution. As noted, the Pueblo linguistic area includes languages with some of the largest inventories of suppletive verbs in our sample. This is a region which has had a robust agricultural system and settled cities since at least 500 BC, with the beginnings of this system stretching back at least a millennium before (Woodbury and Zubrow 1979), and which may have been the origin of the domesticated turkey (Munro 2006). The cluster of languages in Central Mexico, which includes many long-agricultural Otomanguean groups (Campbell 2017), is another piece of counterevidence. If the Caucasus proves to hold as much Sg-Pl alternation as our limited sample in Eurasia suggests, this would be another region of long-settled agriculturalists. In much of the world and especially in our sample, a sedentary and agriculture-centered lifestyle is relatively new. The greater presence of Sg-Pl alternation among historical hunter-gatherer groups is most probably a reflection of both our sample (which excludes LOL languages, see Dahl 2015) and the global distribution of linguistic diversity (i.e. denser among recent hunter-gatherer populations).

Another possibility for the strong tendency for this feature's presence in North America is some as-yet unrecognized cultural feature which promotes the development of these verbal pairs. While this is possible, we are at a loss to hypothesize what this cultural trait might be, nor how it might spread across a continent independently of the language contact necessary for forming a linguistic area. Notably, we were unable to find any trace of Sg-Pl alternation in records of Plains Indian Sign Language (Clark 1885; Mallery 1880/1978), a trade language over a large part of interior North America (but see Samarin 1987 for questions about its robustness at the time of European contact). Incredible as it seems across a continent as large and geographically varied as North America, the best explanation we have for the presence of Sg-Pl alternation in this macroarea is a long-standing series of contact

events neighbor-to-neighbor and linguistic area-to-linguistic area. Sg-Pl alternation would thus have spread through the continent one neighbor and one area at a time, perhaps along with the spread of language families carrying this feature (such as Uto-Aztecan), which is particularly prone to being borrowed thanks to its functional properties.

7 Conclusion

We have presented the largest systematic worldwide survey of Sg-Pl alternation to date, with a very dense sampling in the Americas. Our large sample coupled with fine-grained collected data have enabled us to confirm quantitatively several general observations from previous literature, and have allowed us to distinguish between universal tendencies and more specific areal patterns of Sg-Pl alternation.

The universal factors shaping Sg-Pl alternation affect the typical semantics of verb pairs, largely Motion Events (Talmy 1972), and target the Figure of such events, the syntactic consequence of which is an absolutive alignment. We have proposed a functional explanation for these universal factors, which are related to visual processing of entities in space (see also François 2019). Sg-Pl alternation is also influenced by local factors targeting semantically specific verbs that seem to spread across close neighboring languages. This areal spread does not typically operate at the level of direct borrowing of linguistic forms, but rather through borrowing the concept of lexical differentiation for specific events.

Although most common among North American languages, Sg-Pl alternation is fairly frequent around the world, and appears to be an areal property of at least the Guaporé-Mamoré region and Eastern Papua. Because of the highly local nature of the spread of this feature, it is likely that there are significant areal patterns yet to be uncovered outside of the Americas, where Sg-Pl alternation is less studied. In most languages, only a handful of verbs alternate according to number, and that number distinction overwhelmingly contrasts Singular-Plural, although it is plausible that Individual-Collective semantics are underrepresented in existing grammatical descriptions.

The distribution of Sg-Pl alternation among the world's languages is complex, and our detailed investigation has uncovered global, areal, and highly local factors involved in its development. The multifaceted nature of this phenomenon also highlights the importance of combining quantity (the number of languages surveyed) with quality (the collection of highly detailed information for each language) in typological and areal studies.

A Tripartite number semantics

Tripartite number systems occur in 16 languages in our sample (see Table 17).[33] The Singular-Dual-Plural distinction is by far the most frequent, with only Klamath-Madoc displaying a Singular-Paucal-Plural distinction. As mentioned in Section 4.3, tripartite systems tend to coexist with bipartite ones. They are predominant over bipartite ones in Creek, Karok, Navajo, North Slavey, Southeastern Pomo and Madi, and the only pattern in Maricopa and Western Keres (contra Booker 1982: 24). Note that cases of predominant tripartite systems almost exclusively occur in North America, with the exception of Madi (South America).

Table 17: Distribution of bipartite versus tripartite alternations.

	Language	# of bipartite alternations	# of tripartite alternations	Percent tripartite
Bipartite alternations predominate	Koasati	34	3	5 %
	Tabo[a]	21	5	19 %
	Klamath-Modoc	8	2	20 %
	Imonda[a]	18	6	25 %
	Choctaw[a]	20	9	31 %
	Northern Paiute	12	6	33 %
	Kunama	6	3	38 %
Tripartite alternations predominate	Madi	9	10	53 %
	North Slavey	5	6	55 %
	Karok	2	3	60 %
	Southeastern Pomo	6	9	60 %
	Creek	5	8	62 %
	Navajo	3	5	63 %
	Western Keres	0	61	100 %
	Maricopa	0	4	100 %
Unknown	Taos Northern Tiwa	?	?	?

[a]Some verb alternations in Choctaw (5), Imonda (1) and Tabo (1) are undetermined for bi- versus tripartite patterning.

33 We have also found two languages which seemingly display a quadripartite system. In Tabo (isolate; Papunesia; Schlatter 2003: 43), the verb 'die' displays root suppletion *hoe* 'die.SG' ~ *ha:la* 'die.PL', but a regular pattern limited to a lexically determined verb group "can also be used to stress even greater numbers", yielding *ha:la**ha***, and *ha:la**hale***. In Choctaw (Broadwell 2006: 135; Heath 1980: 21), there are four forms *bali:li* 'run.PAUC' ~ *čaffa* 'run(away?).PL' ~ *tiła:ya* 'run.away.DUAL' ~ *yiłi:pa* 'run.away.TRIAL', but we cannot exclude that these might be two distinct verb pairs ('run' vs. 'run.away').

Acknowledgments: We would like to thank for their collaborative responses: Jon Arensen, Kevin Baetscher, Anna Bugaeva, Joshua Birchall, Seth Cable, Eric Campbell, Mily Crevels, Matthew Davidson, Donna Gerdts, Florian Lionnet, Patrick Littell, Suzi Oliveira de Lima, Maxwell Gomes Miranda, Simone Mattiola, Kirk Miller, Hank Nater, Brigitte Pakendorf, Corey Roberts, Bonny Sands, Antoinette Schapper, Eva Schultze-Berndt, and Adam Werle. We thank our coding team, which included both research assistants and colleagues: Natalia Chousou-Polydouri, Kellen Parker van Dam, Anna Graff, Selma Hardegger, Nora Muheim, and David Timothy Perrot. We owe special thanks to Kellen Parker van Dam for the maps in Figures 1, 5–9. This work was made possible by the financial support of the Out of Asia SNSF Sinergia project CRSII5_183578.

Data availability: All data tables and associated source information are available at https://osf.io/amfdn/.

References

Beck, David. 2000. Grammatical convergence and the genesis of diversity in the northwest coast sprachbund. *Anthropological Linguistics* 42(2). 147–213.

Bierge, Stefanie Ramos. 2017. *Clause types and transitivity in Wixárika (Huichol): A Uto-Aztecan language*. PhD dissertation, University of Colorado at Boulder.

Boas, Franz. 1898. Summary of the work of the committee in British Columbia. *Report of the British Association for the Advancement of Science* 68. 667–688.

Boas, Franz. 1911a. Tsimshian. In Franz Boas (ed.), *Handbook of American Indian languages: Volume 1*, vol. 40 (Bulletin of American Ethnology), 283–422. Washington: Government Printing Office. Available at: http://www.uaf.edu/anla/item.xml?id=TS886B1911.

Boas, Franz. 1911b. Chinook. In Franz Boas (ed.), *Handbook of American Indian languages: Volume 1*, vol. 40 (Bulletin of American Ethnology), 559–678. Washington: Government Printing Office.

Boas, Franz. 1911c. Introduction. In Franz Boas (ed.), *Handbook of American Indian languages* (Bureau of American Ethnology Bulletin 40), 1–83. Washington, D. C.: Government Printing Office.

Böhm, Gerhard. 1984. *Grammatik der Kunama-Sprache*, vol. 22 (Beiträge zur Afrikanistik). Wien: Afro-Pub, Institut für Afrikanistik.

Booker, Karen M. 1982. Number suppletion in North American Indian languages. *Kansas Working Papers in Linguistics* 7. 15–29.

Broadwell, George Aaron. 2006. *A Choctaw reference grammar* (Studies in the anthropology of North American Indians). Lincoln: University of Nebraska Press.

Campbell, Eric W. 2017. Otomanguean historical linguistics: Past, present, and prospects for the future. *Language and Linguistics Compass* 11(4). e12240.

Campbell, Lyle. 1997. *American Indian languages: The historical linguistics of Native America*, vol. 4 (Oxford Studies in Anthropological Linguistics). Oxford: Oxford University Press.

Clark, W. P. 1885. *The Indian sign language*. Philadelphia: L. R. Hamersly & Co.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Crevels, Mily & Hein van der Voort. 2008. The Guaporé-Mamoré region as a linguistic area. *From linguistic areas to areal linguistics* 90. 151–179.

Dahl, Östen. 2015. *How WEIRD are WALS languages?* Paper presented at MPI-EVA, Leipzig.
Durie, Mark. 1986. The grammaticization of number as a verbal category. *Annual Meeting of the Berkeley Linguistics Society* 12. 355–368.
Eggans, Fred. 1979. Pueblos: Introduction. In Alfonso Ortiz (ed.), *Southwest*, vol. 9 (Handbook of North American Indians), 224–235. Washington, D.C.: Smithsonian Institution.
Everett, Caleb. 2006. *Gestural, perceptual, and conceptual patterns in Karitiana*. Houston, Texas: Houston: Rice University dissertation.
Félix Armendáriz, Rolando Gpe. 2005. *A grammar of River Warihio (Mexico)*. PhD dissertation, Rice University.
Frajzyngier, Zygmunt. 1985. Ergativity, number, and agreement. *Annual Meeting of the Berkeley Linguistics Society* 11. 96–106.
François, Alexandre. 2019. Verbal number in lo–toga and hiw: The emergence of a lexical paradigm. *Transactions of the Philological Society* 117(3). 338–371.
van Gijn, Rik. 2006. *A grammar of Yurakaré*. PhD dissertation, Radboud Universiteit Nijmegen Available at: http://webdoc.ubn.ru.nl/mono/g/gijn_e_van/gramofyu.pdf.
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. 2021. *Glottolog 4.4*. Leipzig. https://doi.org/10.5281/zenodo.4761960 (accessed 5 October 2021).
Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.
Haspelmath, Martin. 2013. Occurrence of nominal plurality (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Zenodo.
Haugen, Jason D. & Michael Everdell. 2015. 'To kill' and 'to die' (and other suppletive verbs) in Uto-Aztecan. *Language Dynamics and Change* 5(2). 227–281.
Heath, Jeffrey. 1980. Choctaw suppletive verbs and derivational morphology. *Kansas Working Papers in Linguistics* 5(2). 1–24.
Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33(2-3). 61–83.
Hill, Jane H. 2019. Areal linguistics and linguistic areas in California. In *The Routledge handbook of North American languages*, 572–586. Abingdon: Routledge.
Hofherr, Patricia Cabredo & Brenda Laca. 2012. *Verbal plurality and distributivity*. Berlin, Boston: De Gruyter.
de Hollenbach, Elena Erickson. 2013. *Grama´tica del mixteco de Magdalena Peñasco (Sa'an Ñuu Savi)*. México, D.F.: Instituto Lingüístico de Verano. 475.
Inman, David. 2021. The exponence and development of plurals in nuuchahnulth. *Papers for the 56th International Conference on Salish and Neighbo(u)ring Languages*, 111–133.
Jacobs, Melville. 1937. Historic perspectives in Indian languages of Oregon and Washington. *The Pacific Northwest Quarterly* 28(1). 55–74.
Jeanne, LaVerne, Kenneth Hale & Paula Pranka. 1984. *Where is suppletion?* Paper presented at GLOW 1984.
Kinkade, M. Dale. 1981. Singular vs. plural roots in Salish. *Anthropological Linguistics* 23(6). 262–269.
Krasnoukhova, Olga. 2022. Number in the languages of South America. In Paolo Acquaviva & Michael Daniel (eds.), *A comparative handbook*, 609–670. Berlin, Boston: De Gruyter Mouton.
Lindenfeld, Jacqueline. 1973. *Yaqui syntax*, vol. 76 (University of California Publications in Linguistics). Berkeley & Los Angeles: University of California Press.
Mallery, Garrick. 1880/1978. A collection of gesture-signs and signals of the North American Indians with some comparisons. In Jean Umiker-Sebeok & Thomas A. Sebeok (eds.), *North America: Classic*

comparative perspectives, vol. 1 (Aboriginal Sign Languages of the Americas and Australia), 77–406. New York: Plenum Press.

Marlett, Stephen A. 1981. *The structure of Seri*. PhD dissertation, University of California at San Diego.

Mattiola, Simone. 2019. *Typology of pluractional constructions in the languages of the world* (Typological Studies in Language Volume 125). Amsterdam & Philadelphia: John Benjamins.

Mattiola, Simone. 2020. Pluractionality: A cross-linguistic perspective. *Language and Linguistics Compass* 14(3). https://doi.org/10.1111/lnc3.12366.

Miller, Wick R. 1965. *Acoma grammar and texts*, vol. 40 (University of California Publications in Linguistics). Berkeley & Los Angeles: University of California Press.

Miranda, Maxwell Gomes. 2014. *Morfologia e morfossintaxe da li´ngua Krahô (família Jê, tronco Macro-Jê)* [Morphology and morphosyntax of the Krahô language (Jê family, Macro-Jê).]. PhD dissertation, Universidade de Brasília.

Mithun, Marianne. 1988. Lexical categories and the evolution of number marking. In Michael Hammond (ed.), *Theoretical morphology: Approaches in modern linguistics*, 211–234. Nueva York: Academic Press.

Mithun, Marianne. 2022. Sitting and talking together: Packaging meaning into verbs with the neighbors. *Linguistic Typology* 26(2). 375–402.

Müller, Neele. 2014. Tense-aspect-mood-evidentiality (TAME). In Pieter Muysken, Harald Hammarström, Olga Krasnoukhova, Neele Müller, Joshua Birchall, Simon van de Kerke, Loretta O'Connor, Swintha Danielsen, Rik van Gijn & George Saad (eds.), *South American Indian language structures (SAILS) online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at: https://sails.clld.org.

Munro, Natalie D. 2006. The role of the Turkey in the Southwest. In Douglas H. Ubelaker, Dennis J. Stanford, Bruce D. Smith & J. E. Szathmáry (eds.), *Environment, origins, and population*, vol. 3 (Handbook of North American Indians), 463–470. Washington, D.C.: Smithsonian Institution.

Murane, Elizabeth. 1974. *Daga grammar: From morpheme to discourse*. Microfiche ed. vol. 43 (Summer Institute of Linguistics: Publications in Linguistics). Glendale, California: Dallas, Texas: The Summer Institute of Linguistics and the University of Texas at Arlington.

Nakagawa, Hiroshi. 2022. Verbal number. In Anna Bugaeva (ed.), *Handbook of the Ainu language*, 573–612. De Gruyter.

Nakayama, Toshihide. 2001. *Nuuchahnulth (nootka) morphosyntax*, vol. 134 (University of California Publications in Linguistics). Includes bibliographical references, 159–165. Berkeley & Los Angeles: University of California Press.

Nellis, Jane G. & Neil Nellis. 1983. *Diccionario zapoteco de Juárez: zapoteco-español, español-zapoteco (Ca titsa' qui' ri'u)* [Zapotec Dictionary of Juárez: Zapotec-Spanish, Spanish-Zapotec (Ca titsa' qui' ri'u)], vol. 27 (Serie de vocabularios y diccionarios indígenas "Mariano Silva y Aceves"). Mexico: Instituto Lingüístico de Verano.

Newman, John. 2002. *The linguistics of sitting, standing, and lying*, vol. 51 (Typological Studies in Language). Amsterdam: John Benjamins.

Pebesma, Edzer. 2018. Simple features for R: Standardized support for spatial vector data. *The R Journal* 10(1). 439–446.

Pebesma, Edzer & Roger Bivand. 2023. *Spatial data science: With applications in R*. Boca Raton: Chapman and Hall/CRC. Available at: https://r-spatial.org/book/.

Rzymski, Christoph, Tiago Tresoldi, Simon J. Greenhill, Mei-Shin Wu, Nathanael E. Schweikhard, Maria Koptjevskaja-Tamm, Volker Gast, … Johann-Mattis List. 2020. The database of cross-linguistic colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific Data* 7(1) 1–12.

Samarin, William J. 1987. Demythologizing plains Indian sign language history. *International Journal of American Linguistics* 53(1). 65–73.

Sapir, Edward & Morris Swadesh. 1955. *Native accounts of Nootka ethnography*. Bloomington: Indiana University, Research Center in Anthropology, Folklore, and Linguistics.

Saxton, Dean, Lucille Saxton & Susie Enos. 1983. *Dictionary Tohono Oodham/Pima to English*, 2nd edn. Tucson: University of Arizona.

Schlatter, Tim. 2003. *Tabo language grammar sketch (Aramia River)*. Ms.

Sherzer, Joel. 1976. Areal linguistics in North America. In Thomas A. Sebeok (ed.), *Native languages of the Americas: Volume 1*, 121–174. New York: Plenum.

Shibatani, Masayoshi. 1999. *The languages of Japan* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

Smith, Bruce D. 2006. Plant and animal resources: Introduction. In Douglas H. Ubelaker, Dennis J. Stanford, Bruce D. Smith & J. E. Szathmáry (eds.), *Environment, origins, and population*, vol. 3 (Handbook of North American Indians), 219–222. Washington, D.C.: Smithsonian Institution.

Storto, Luciana. 2014. Reduplication in Karitiana (Tupi). In *Reduplication in indigenous languages of South America*, 401–426. Leiden, The Netherlands: Brill.

Suárez, Jorge A. 1983. *La lengua Tlapaneca de Malinaltepec* [The Tlapanec language of Malinaltepec.]. México: Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México.

Talmy, Leonard. 1972. *Semantic structures in English and Atsugewi*. PhD dissertation, University of California, Berkeley.

Talmy, Leonard. 2000. *Toward a cognitive semantics* (Bradford book v. 2). London: MIT Press.

Tarpent, Marie-Lucie. 1983. Morphophonemics of nisgha plural formation: A step towards proto-tsimshian reconstruction. *Kansas Working Papers in Linguistics* 8(2). 123–214.

Thompson, Laurence C. & Marvin Dale Kinkade. 1990. Languages. In Wayne Suttles (ed.), *Northwest coast*, vol. 7 (Handbook of North American Indians), 30–51. Washington, D.C.: Smithsonian Institution.

Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic analyses: The non-Bantu languages of North-Eastern Africa* (Handbook of African Languages), 2nd edn. London: Oxford University Press.

Veerman-Leichsenring, Annette W. 1991. *Grama´tica del Popoloca de Metzontla. Spanish and Popoloca Summary in English Includes bibliographical references*, 544–550. PhD dissertation, Rijksuniversiteit te Leiden.

Veselinova, Ljuba N. 2003. *Suppletion in verb paradigms: Bits and pieces of a puzzle*. PhD dissertation, Stockholm University.

Veselinova, Ljuba N. 2006. *Suppletion in verb paradigms: Bits and pieces of the puzzle* (Typological studies in language v. 67). OCLC: ocm69680110. Amsterdam & Philadelphia: John Benjamins.

Veselinova, Ljuba N. 2013. Verbal number and suppletion (v2020.3). In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Zenodo. https://10.5281/zenodo.13950591.

Wickham, Hadley. 2016. *Ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. Available at: https://ggplot2.tidyverse.org.

Willett, Thomas L. 1991. *A reference grammar of Southeastern Tepehuan*. vol. 100 (Summer Institute of Linguistics: Publications in Linguistics). Dallas, Texas: The Summer Institute of Linguistics and the University of Texas at Arlington.

Woodbury, Richard B. & Ezra B. W. Zubrow. 1979. Agricultural beginnings, 2000 BC–AD 500. *Handbook of North American Indians* 9. 43–60.