Areal typology

Peter Arkadiev, University of Potsdam

Submitted to:
Handbook of Historical Sociolinguistics, edited by Bridget Drinka, Terttu Nevalainen & Gijsbert Rutten. De Gruyter Mouton.
Revised version, August 2026

Abstract: Areal typology studies the geographical distribution of languages and linguistic features, thus lying at the intersection of language typology, historical linguistics and the study of language contact. This chapter discusses some of the prominent ways in which geographical considerations are relevant for typological studies, such as the notion of "macroarea" and its role in language sampling, or the distinction between "spread zones" and "residual/accretion zones". Several recent approaches in areal typology, both qualitative and quantitative, are surveyed. After critically examining the conceptual and methodological problems associated with identification and delimitation of "linguistic areas", it is suggested that such notions as "contact superposition zone" and "areal cline" might be more adequate. In general, the chapter advocates a deeper integration of typological, comparative-historical, contact-linguistic and sociolinguistic approaches to linguistic diversity on a broader interdisciplinary background.
(12508 words)

1. Introduction

Areal typology (or areal linguistics; I consider these terms largely synonymous) is the study of the geographical distribution of linguistic features and the historical processes giving rise to these distributions. Thus areal typology lies at the intersection of language typology, historical linguistics and the study of language contact and is informed by methods and results of these three domains. As the name "areal typology" suggests, this discipline can be considered an integral part of linguistic typology as a study of linguistic diversity in general (Velupillai 2012: 15–16). This is especially true in the light of the observable shift of focus in typology during the last three decades from the quest for putative language universals towards questions of the type that Bickel (2007) called "what's where why", cf. also Bickel (2017). In this sense almost any typological study going beyond simple enumeration of attested types of phenomena becomes to a certain extent a study in areal typology, since its empirical results by necessity have a geographical dimension in the form of observations of the kind "type A is particularly well-attested in Australia", "type B is almost non-attested in Eurasia" or "type C is more or less randomly distributed across the globe". Testing the validity of such observations and accounting for them, also in historical terms, is one of the important objectives of language typology as an explanatory discipline. Here recourse to insights and methods of historical and contact linguistics, as well as more generally of human history and related fields, becomes inevitable. Besides that, areal considerations play an important methodological role in typological studies based on language sampling, insofar as worldwide language samples have to be both representative and balanced with respect to the geographical distribution of language families in the world (Bakker 2010).

The goal of this article is to summarise the most important notions and trends in areal typology understood as outlined above. I shall start by discussing in somewhat greater detail the major ways in which (linguistic) geography is relevant for linguistic typology (section 2) and then shall dwell upon the traditional understanding of the notion of "linguistic area", its pitfalls and potential improvements (section 3), and finally turn to ways these and related problems have been addressed in more recent (areal) typological work (section 4).

2. Geographical dimensions of linguistic typology

2.1. Macroareas vs. microareas

As said above, once typology goes beyond either pure taxonomy of linguistic phenomena or (often speculative) hypotheses aiming at explaining the attested vs. non-attested phenomena in terms of some functional or structural principles, it almost inevitably assumes a geographical dimension. A perfect illustration of this point is the *World Atlas of Language Structures* (WALS, Haspelmath et al. eds. 2005, Dryer & Haspelmath eds. 2013), whose chapters, on the one hand, contain information about the proposed classificatory values for each of the studied structural features (more than 190 in total) and, on the other, plot the geographical distribution of the languages showing each of the values thus defined. While the chapters of WALS themselves can serve as succinct reference overviews on the classifications of selected phenomena, the most important value of WALS is constituted precisely by its having enabled linguists for the first time to visualise the geographical distribution of a considerable number of variables of linguistic diversity ranging from phonology to the lexicon and to study potential interactions between these variables and the geographical distributions of their combined values (specifically on this aspect of WALS see e.g. Comrie 2020; it is also worth mentioning a number of regional atlases inspired by WALS, such as *The Areal Typology of Languages of the Americas*, Inman et al. 2025, *Atlas of the Balkan Linguistic Area*, Adamou & Sobolev 2025, *Typological Atlas of the Languages of Daghestan*, Moroz et al. 2026).

A cursory look at more or less randomly selected maps from WALS is enough to get an idea that geography may play an important role in the distribution of typological features and, at the same time, that the spatial distributions of different phenomena may be drastically dissimilar from each other. For example, the feature 87A "The order of adjective and noun" (Dryer 2013), Figure 1, shows a geographical distribution that can be described in terms of subcontinent-size regions, with e.g. the whole of Northern Eurasia representing the "adjective-noun" order, while most of Africa, apart from "pockets" in Ethiopia and Sahel, as well as parts of Western Europe and the Middle East representing the "noun-adjective" order. Also notable is the fact that even if we zoom in on this map, we apparently do not find many regions where languages with different orders of adjective and noun would be interspersed randomly. Such clear and large-scale geographical distributions, which are also observed for a number of other typological variables, can hardly be accidental and require an explanation.

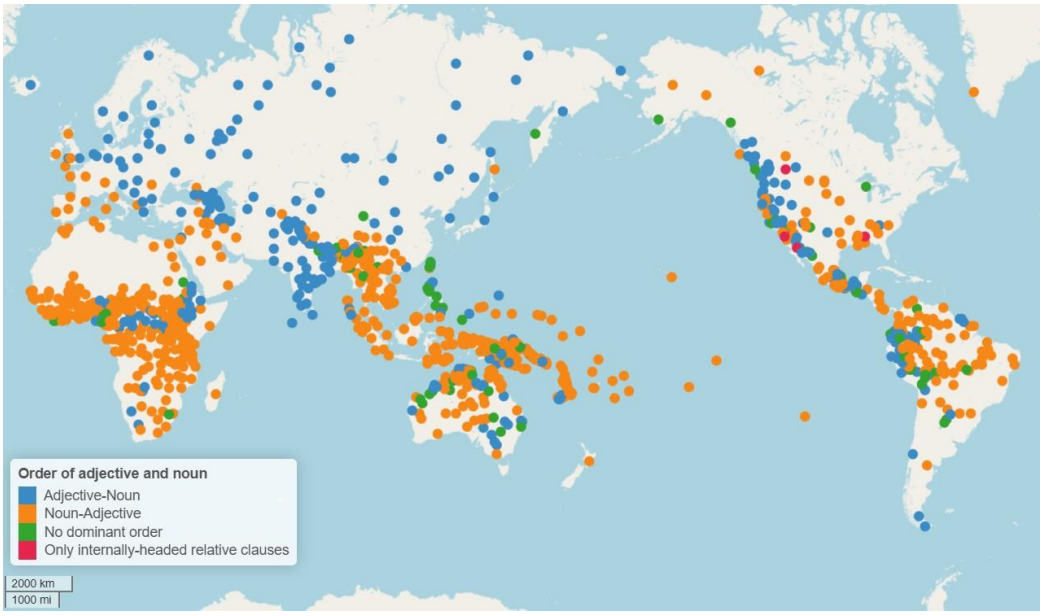

Figure 1. Order of adjective and noun based on the data from WALS (Dryer 2013), created in R with the lingtypology package (Moroz 2017)

If we take a different feature, e.g. 19A "Presence of uncommon consonants" (Maddieson 2013), Figure 2, we also observe some clearly non-random geographical patterns, but of a different kind. Perhaps unsurprisingly, there are no more continent-size distributions; by contrast, against the background of vastly predominating languages with "dull" consonant inventories lacking any rare segment types, we find compact geographical regions where languages with a particular type of uncommon consonants are concentrated. Such regions are, for instance, the so-called Macro-Sudan Belt in Africa stretching from the Guinean coast to the border between the Congo basin and the East African Highlands, where languages with labiovelar consonants are found, or the Caucasus, where languages with pharyngeal consonants are well-attested. Even if apparently less spectacular than the geographical distributions of word-order features, such spatial patternings of linguistic properties are likewise non-random and need explanations.

Figure 2. Presence of uncommon consonants based on the data from WALS (Maddieson 2013), created in R with the lingtypology package (Moroz 2017).

The two examples just given illustrate the two major ways in which the geographical dimension enters typological research, i.e. "macroareal" and "microareal". Macroareas and microareas are defined by different criteria, serve different purposes and are studied by different methods. Macroareas are very large (continent-size) geographical regions populated by hundreds or even thousands of languages showing significant genealogical and typological diversity, which is supposed to have emerged and developed independently of that in other macroareas. By contrast, microareas are compact geographical regions with a relatively circumscribed number of languages resp. limited genealogical diversity, and these languages show some structural similarities to each other, usually assumed to be due not (only) to common heritage, but (primarily) to contact-induced diffusion. Unsurprisingly, definitions of microareas crucially rely on linguistic properties (see section 3). It is thus expected that microareas should form a complex structure, and that a smaller microarea may itself form part of one or more larger microareas (cf. Hammarström & Donohue 2014: 169; Wiemer 2019; Wiemer & Arkadiev 2025). The most important differences between macro- and microareas are summarised in Table 1.

Table 1. Macroareas vs. microareas

	macroareas	**microareas**
geographical extent	large (continent-size)	compact
definitional criteria	primarily geographic	primarily linguistic
genealogical diversity	large	limited
typological diversity	large	limited
temporal depth	large (human settlement)	small (historical periods)
main factors	large-scale population movements	small-scale human interactions and contacts
purpose of study	control in language sampling and/or statistical modeling, speculations about linguistic prehistory	study of diffusion of linguistic features in contact situations

Two important caveats are in order. First, one should distinguish between geographical distributions of individual linguistic traits, as exemplified by WALS maps above, and "linguistic areas" usually defined either as abstractions over multiple such distributions (see section 3) or without regard to them at all, on purely geographical basis (the latter particularly pertains to macroareas). Second, the distinction between macroareas and microareas is gradual rather than discrete. Indeed, if macroareas are to be defined in terms of continents, any area smaller than that is automatically assigned to the class of "microareas", which may concomitantly include areas with considerable geographical extent, historical depth and linguistic diversity. A case in point is the aforementioned Macro-Sudan Belt in Africa, which Güldemann (2008) argues to constitute a large and ancient region of multilateral language contact leading to diffusion of a whole array of linguistic features. See also Nichols (this volume and references therein) on such "smaller macroareas".

2.2. Macroareas and language sampling

The initial purpose of considering macroareas in typological work has been to use them "as (part of a scheme for) control for areal and genealogical effects" in language sampling (Hammarström & Donohue 2014: 169), since it is natural to assume that the (linguistic) populations of continents separated by large bodies of water developed largely independently during much of (recent) human history[1]. Therefore, macroareas can be defined "based purely on principles of geographical independence" (ibid.: 169), and moreover, as these authors argue, must be thus defined in order for the languages and genealogical groupings in different macroareas to be considered independent. The first influential proposal regarding macroareas was made by Dryer (1989), where he divided the world into five macroareas defined primarily (although not exclusively) by large geographical divisions, i.e. Africa, Eurasia, Australia and New Guinea, North America and South America. Dryer used macroareas in order to test putative implicational universals in the domain of word order: a certain frequency distribution of typological variables (which Dryer counted in terms of genera rather than languages) can be considered universally significant only if it is observed in all macroareas. This can be illustrated by different outcomes of testing for association between the mutual orders of demonstrative and noun vs. adjective and noun (Dryer 1989: 272): while noun-demonstrative order is strongly correlated with noun-adjective order in all macroareas (Table 2), the opposite demonstrative-noun order allows both adjective-noun and noun-adjective order with different proportions in different macroareas (Table 3). (Larger numbers are highlighted in both tables.)

[1] A notable exception to this is the interaction between northeastern Africa and southwestern Eurasia, involving population and language movements spanning several millennia, see e.g. Güldemann (2018: 509–510).

Table 2. Noun-demonstrative order and mutual order of adjective and noun

	Africa	Eurasia	Australia and New Guinea	N.America	S.America	total
NDem&NAdj	**28**	**14**	**8**	**8**	**5**	63
NDem&AdjN	1	2	0	1	0	4

Table 3. Demonstrative-noun order and mutual order of adjective and noun

	Africa	Eurasia	Australia and New Guinea	N.America	S.America	total
DemN&NAdj	5	**27**	**7**	**19**	**8**	66
DemN&AdjN	**7**	10	6	14	6	43

Since Dryer's initial insights, a number of refinements have been proposed to language sampling procedures taking macroareas as their basis, e.g. the Genus-Macroarea method (Miestamo 2005, Miestamo et al. 2016) designed for so-called variety sampling (see Rijkhoff et al. 1993 on the distinction between probability samples and variety samples), whose aim is to capture the world's linguistic diversity in the most representative way. Besides that, methods for circumventing possible biases by post-sampling statistical controls have been proposed, see e.g. Bickel (2013) and Guzmán Naranjo & Becker (2022).

Dryer's method can be, *mutatis mutandis,* applied in order to both test for putative universals (or lack thereof) and to discover (macro)areal patterning of virtually any linguistic phenomenon. The latter goal has been pursued in particular by Johanna Nichols and her associates within the so-called "Predictive Areality Theory", see e.g. Nichols (1992, 1993, 2017), Nichols et al. (2004), Bickel & Nichols (2006), Bickel (2020), Hartmann & Nichols 2025. Starting with already mentioned areas of intermediate size, Nichols also considers macroareas larger than continents (e.g. "Old World" = Africa + Eurasia vs. "New World" = the Americas") or defined independently of continents, e.g. the Pacific Rim, i.e. the area around the Pacific, including Australia, New Guinea and the islands to the east of the Wallace line, as well the coastal areas of Eurasia and the Americas. The historical depth of such huge areas can go up to 60 000 years, and the genealogical and typological diversity of the languages populating them is huge. Therefore, unsurprisingly, one does not expect such areas to show homogeneity or exclusivity with respect to any linguistic feature (cf. Bickel & Nichols 2006 on "variance" and "leakage"); instead, what one does observe, similarly to Dryer's results, are skewed frequency distributions. For instance, Africa and the Pacific Rim show a higher concentration of languages with possessive classification (Nichols & Bickel 2013) than the rest of the world, see Figure 3; likewise, as Nichols (1993: 39) shows, "Australia and Eurasia have far more ergative languages than expected, the New World has far more stative-active languages, and Africa has far fewer non-accusative languages".

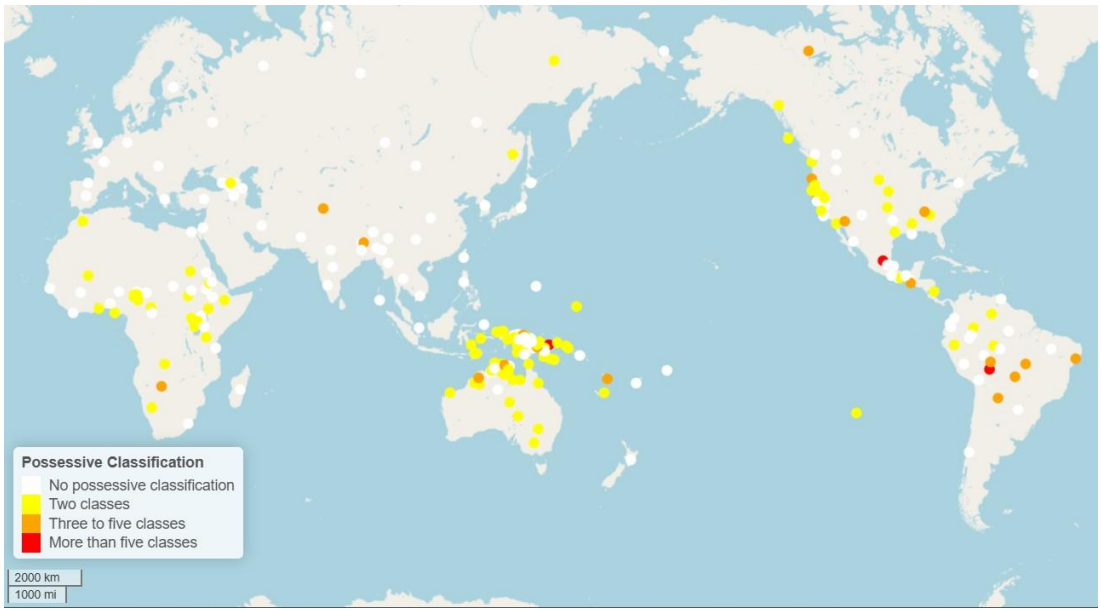

Figure 3. Possessive classification based on the data from WALS (Nichols & Bickel 2013) , created in R with the lingtypology package (Moroz 2017).

Of particular interest are also macro-areal clines (on the notion of areal cline see in more detail in section 3) involving gradual, rather than discrete, typological features showing an increase resp. decrease spanning across large territories. As an example consider the feature of finiteness in clause-combining (Shagal et al. submitted), defined as the proportion of subordinate clause types employing finite verbal forms (i.e. encoded in the same way as verbal forms in independent declarative clauses). According to the authors, this feature shows a robust west-to-east distribution across Northern Eurasia, with languages of Europe on average featuring the highest and the languages of East Asia the lowest proportion of finite subordination. A similar west-east distribution is reported for some other gradual features as well, e.g. for the inflectional encoding of person (Nichols 2017); a recent study of phonological complexity (Hartmann & Nichols 2025) likewise reports a variety of large-scale clines. Again, and importantly, these are trends whose validity hinges on sampling procedures, definitions of comparative concepts, data encoding, and statistical methods applied.

2.3. Problems with macroareas

Under the assumption that macroareal trends are empirically valid, how can they be explained? While diffusion through language contact has always been invoked as an explanation (and often as a definitional feature) of linguistic microareas (see section 3), contact alone can hardly account for distributions with possible time-depth of millennia and spanning across thousands of miles let alone involving both sides of the Pacific. Explanatory factors inevitably take recourse to large-scale population movements in pre-historic times. For instance, Bickel & Nichols (2006: 8) write that

> "the P[acific]R[im] formed as coastally adapted people, and their languages and cultures, spread out of Southeast Asia beginning late in the last glaciation and continuing into recent centuries with the Austronesian spread and the Chukchi spread to the Bering Strait. They spread coastally, as is shown by the striking coastal distributions of variables such as V-S order and multiple possessive classes."

Such and similar proposals are, in my view, fascinating and thought-provoking speculations about the past so distant and often so obscure that we can only hope that

collaboration between linguists, archaeologists, anthropologists and population genetics would be able to devise methods able to test them (cf. e.g. Heggarty 2017 on possible directions for such a synthesis).

Another problematic aspect of the work on macroareal typological distributions is purely linguistic and concerns the choice of variables and features mapped. While Bickel & Nichols (2006: 2) in their critique of feature-based work on areal typology rightly point out that "variables exhibiting the requisite isoglossic behavior may have to be defined as an abstraction which is in itself unlikely to be able to diffuse", the very same, in my view, applies to some of the features or feature aggregates invoked in their own work. Consider, for example, ergativity discussed in Nichols (1993). While a widely accepted typological feature, ergativity is in fact an abstraction over distinct manifestations, such as locus of expression (head-marking vs. dependent-marking), and restrictions related to nominal subcategories, tense and aspect, clause-types, lexical verb classes, discourse pragmatics etc. When taken into account (in full accordance with Bickel and Nichols' own "autotypologising" multivariate approach, see e.g. Bickel & Nichols 2002, Witzlack-Makarevich & Bickel 2019), such finer-grained variables almost inevitably lead to purported macroareal trends dissipating into a variety of much more local patterns of cross-linguistic similarity. The latter then more readily lend themselves to explanations in terms of better-understood mechanisms of genealogical inheritance and language contact.

2.4. Further classifications of areas

Turning from macroareas to areas of lesser size, another highly important way in which geography is relevant for linguistic typology has also been emphasised in the work by Nichols (1992, 1997, 2015, 2020), where it has been argued that distributions of typological diversity across geographically defined regions correlate with the historically attested or reconstructed patterns of dispersal of human resp. linguistic populations in these regions, which, in turn, are inevitably constrained by geography itself (cf. also Huisman et al. 2019). Nichols (1992: 13–24) introduced the now famous distinction between "spread zones" and "residual zones", the latter being renamed into "accretion zones" in subsequent work (Nichols 1997: 369). This dichotomy, which Nichols herself (1992: 13) did not intend as an exhaustive classification, is based on the genealogical diversity of the area and its development in time. Spread zones show relatively low degrees of diversity, being often dominated by a single language family, and such dominant language families, at least in principle, can replace one another over time, with earlier language groupings becoming extinct or pushed to the margins, sometimes surviving as isolates (see e.g. Urban 2021). Examples include central Eurasia, where Iranian, Turkic and Mongolic languages spread one after another in historical times largely displacing the original populations, or central and southern Australia, where several branches of the Pama-Nyungan family underwent expansion during the last millennia (McConvell 1996). In geographical terms, spread zones are often vast plains lacking major natural obstacles to population expansion. By contrast, residual, or accretion zones, are characterised by higher degrees of genealogical diversity, which, besides, increases over time due to regular immigration of newcomers and does not show any dominant family or language. Examples include the Caucasus, discussed in detail by Nichols in many of her works, California and New Guinea. In terms of their geography, residual zones are often mountainous or otherwise hardly accessible regions. Table 4, based on Nichols (1992: 16–17; 21), summarises the main characteristic properties of the two types of zones.

Table 4. Spread zones vs. accretion zones

	spread zones	**accretion zones**
genealogical diversity	low	high
typological diversity	low	high
time-depth of language families	low	high
effect of time	spread and succession of languages and lineages	accumulation of languages and lineages
inner structure	centre vs. periphery	no clear centre of innovation
change of diversity in time	no increase	increase
pattern of multilingualism	dominant language as lingua franca	local multilingualism, no lingua franca

In later work (Nichols 2013, 2015, 2020), a more refined classification has emerged with several further dimensions added to the initial dichotomy, including the "vertical" one (see also Urban 2020). According to Nichols, (ethno)linguistic areas localised around mountain ranges have an "uphill direction of language spread which leaves truncated branches, isolates, and archaisms in the highlands and brings in innovations and new languages from the lowlands" (Nichols 2015: 270), as demonstrated by such regions as the Caucasus or the Hindukush (Liljegren 2020). By contrast, plateaus surrounded by mountains, which Nichols (2015) calls "altiplano" and exemplified by the Andean Altiplano itself (on which see Urban 2019) or by Tibet, show a completely different profile. Since such a region is difficult to access and ecologically challenging, "[o]nce a society (with its economy, culture, and language) has adapted to such an area and spread over it, it is difficult to dislodge" and, therefore, "the succession of spreading languages comes not from the outside but from inside" (Nichols 2015: 271), i.e. from dialects or daughter languages. While in mountain crest accretion zones the natural direction of language spread is from lowlands uphill, an altiplano language can spread downhill. Altiplanos thus illustrate what Nichols calls "closed spread zones", i.e. "spread zone surrounded by ecological barriers and often with a distinctive ecology requiring special adaptation" (Nichols 2020: 26), as opposed to "open spread zones", "where entry is relatively unhindered and successive spreads are likely to involve different languages" (Nichols 2015: 274). Closed spread zones can be geographically set off not only by mountains, but also by large bodies of water, an example being the whole of Australia with its highly distinctive typological profile. According to Nichols (2015: 280), closed spread zones "behave much like bottlenecks" in that "their long histories of extinction and convergence reduce the number of descendants relative to the number of ancestors and the number of types", and at the same time being able to "produce extreme development and elaboration of unusual typological variables". Finally, in Nichols (2020) a further type of area is added to the (still non-exhaustive) classification, i.e. "crossroads", which have "a fairly diverse language population including a number of large languages spoken in complex societies" as well as "a history of multiple influences flowing in, out, and through the area but no area-wide language spreads" (ibid.: 30), with examples such as Mainland Southeast Asia or the Balkans.

Nichols' proposals, in particular, the distinction between spread and residual zones and its application in her 1992 book, have been criticised, see e.g. Campbell (2002: 55–58), who called it "a misleading idealization". It is difficult to disagree with Campbell regarding "idealization", however, dismissing it entirely as "misleading" would rather amount to throwing the baby out with the bathwater. While it is clear that every single geographical region has its own unique history of population movements, sociocultural developments and language spreads, a quest for recurrent trends and systematic distinctions is still a worthwhile endeavour, and the very idea that areal typology should be firmly grounded in geography and human history

can hardly be disputed. This is particularly true in the light of impasses of traditional areal linguistics, to which I turn in the next section.

3. "Linguistic areas" and what to replace them with

The reader might wonder why the definition of "linguistic area", apparently the main object of study of areal typology, has not been provided and discussed at the very beginning of this chapter, even more so given that the preceding section by no means avoided using this term. The reason for this compositional "quirk" — quite deliberate, I must admit — is manifold. First, having introduced the important distinction, or rather continuum, between "macroareas" and "microareas", I pointed out that, paradoxical though it may seem, these notions cannot be regarded as denoting two different kinds of the same thing (cf. Muysken 2008: 4). This means that "linguistic area", in the way the term is usually understood, is in fact neither an overarching nor a basic notion of areal typology. Rather, it roughly corresponds to "microarea", with the caveat that "micro-" can mean anything up to a subcontinent in size. Second, in the same passage I hinted at the most important, in fact, definitional characteristic usually associated with linguistic areas, i.e. that the languages populating an area must show structural similarities and these similarities must be due rather to contact than to common descent. As an attentive reader should have noticed, however, this idea, which I shall elaborate upon immediately below, has not played a really prominent role in the discussion of macroareal research or Johanna Nichols' classification of linguistic areas. Third, as is clear already from the title of this section, I am highly skeptical about the very validity of the notion "linguistic area" in its traditional understanding and doubt that it can be useful for (areal) typology unless its conceptual and methodological underpinnings are thoroughly reconsidered. In this skepticism, I am neither alone nor unprecedented, see e.g. Dahl (2001), Masica (2001), Stolz (2002), Enfield (2005), Bisang (2006, 2010), Campbell (2006, 2017), Muysken (2008), among many others starting as early as Weinreich (1958: 378–379), and therefore below I shall only briefly outline the history of the term "linguistic area" and the difficulties and pitfalls associated with it. For more comprehensive recent overviews, see van Gijn & Wahlström (2023) and Wiemer & Arkadiev (2025).

3.1. How to define a "linguistic area"?

The term "linguistic area" is the English version of the German term *Sprachbund* (also often used in English, see e.g. Mišeska Tomić 2008), itself a translation of the Russian term *jazykovoj sojuz* 'language union' proposed by Nikolai Trubetzkoy (1923, 1930)[2,3]. Before, or parallel to, Trubetzkoy, the idea of grouping languages on the basis of convergence rather than genealogy has become prominent in Balkan studies (Sandfeld 1930, see Joseph & Friedman, this volume), as well as in the work of Franz Boas (1917, 1929 etc.) in the United States, who saw linguistic convergence as but one aspect of cultural diffusion in general (Boas 1937). The term Sprachbund has been taken up and further developed by Trubetzkoy's close associate Roman Jakobson (1931, 1972[1938]), who, however, did not give it much prominence in his later work. The term "linguistic area" became current after the publication of Emeneau (1956), who himself borrowed it from an earlier article by Velten (1943: 271), who in turn acknowledged it to be a translation of Trubetzkoy's *Sprachbund*. A further milestone publication is the monograph Masica (1976), where the term "areal typology" is introduced (p. xi) and methodological issues

[2] Interestingly, the term "linguistic area" has been subsequently introduced into both German (*Sprachareal*) and Russian (*jazykovoj areal*) linguistic traditions, where the respective terms have apparently been lately gaining more currency.

[3] It is worth noting that the notion *Sprachbund* as conceived by Trubetzkoy was not free of certain ideological overtones that can be considered imperialistic if not chauvinistic, see e.g. Toman (1995: 185–215), Van Pottelberge (2001), Nekula (2003).

are seriously addressed. After these foundational works, many linguistic areas have been proposed: beside the Balkans and South Asia, those are, for instance, Ethiopia (Ferguson 1970; Tosco 2000), Caucasus (Tuite 1999; Chirikba 2008), Arnhem Land in Northern Australia (Heath 1978), Western Amazonia (Aikhenvald 2002), Mesoamerica (Campbell et al. 1986) etc., see the overview chapters in Hickey (ed.) (2017). Particular attention has been paid to the study of Europe as a linguistic area (for a critique, see Van Pottelberge 2001), cf. Haspelmath (2001) and the series of publications by the research programme EUROTYP in the 1990s to early 2000s, e.g. van der Auwera (ed.) (1998), Dahl (ed.) (2000), Plank (ed.) (2003), as well as some subsequent work, e.g. Heine & Kuteva (2006), Stolz et al. (2006, 2008, 2017), Cysouw (2011), and Drinka (2017), see Grzega (ed.) 2025 for the most recent summary.

Many definitions have been proposed for the term "linguistic area", not all of them equivalent (see e.g. Campbell 2006 for an overview). The following quotation from Enfield (2005: 190) highlights the most important characteristics included in most definitions.

> "A linguistic area is defined as a geographical region in which neighboring languages belonging to different language families show a significant set of structural properties in common, where the commonality in structure is due to contact and where the shared structural properties are not found in languages immediately outside the area (ideally where these include languages belonging to the same families as those spoken inside the area)."

Such and similar definitions of linguistic areas, however, in fact raise more questions than they answer, as evidenced by the immediately following text from Enfield (2005: 190–191):

> "At what level do we say that different languages belong to different families? Is it enough that they are of different subgroups? What degree of structural parallelism counts as significant? How many properties need to be shared? Which ones count? Can we rank their importance? Should typologically unusual features take priority when typologically common features may just as well be diffused? How significant is it that two languages share a handful of features, given thousands of differences?"

Indeed, examination of virtually any linguistic area proposed in the literature reveals a similar bundle of problems arising when one tries to operationalise the definitions of the concept, most of which are variations along the lines of the quotation from Enfield (2005) above. Any "linguistic area" thus defined presupposes at least three essential components: (i) a set of languages *L*; (ii) a set of structural properties *F* they have in common; (iii) a geographical region *G* with certain boundaries that these languages occupy. Now what if different members of *L* share different subsets of the features from *F*? (As is virtually always the case, which has given rise to concepts of "gradual membership", "core" vs. "periphery" and the use of isopleths rather than isoglosses, see e.g. van der Auwera 1998, Dahl 2001: 1458). What if a language located within the boundaries of *G* does not share (enough of) the features from F*?* (An example could be Basque against the neighbouring "Standard Average European" languages.) What if, by contrast, a language located outside of *G* (in the ultimate case, not even sharing boundaries with *G*) shares enough of the features from *F*? (An example could be Southwestern Ukrainian with respect to the Balkan Sprachbund, see Danylenko 2019; cf. Campbell 2017: 28 on "trait-sprawl areas".) What if, in general, at least some of the features from *F* are attested in a much broader geographical region properly including *G*? (As is the case with a whole number of so-called "Balkanisms", which are found across Europe and/or Mediterranean.) Are manifestations of a given feature in different languages, in particular if they are genealogically unrelated, similar enough to each other to warrant cross-linguistic identification? (This is, of course, a

more general methodological problem in typology, see e.g. Stassen 2010.) Ultimately, some kind of inherent circularity appears to be unavoidable in the delimitation of traditional linguistic areas: If we start with some a priori selected set of languages which we believe to constitute a Sprachbund, we unavoidably end up "cherry-picking" structural properties in such a way that they identify this exact set of languages. However, given a worldwide distribution of languages together with their numerous and variegated structural properties, how can we be sure that the features selected in this way are not accidental against the background of all other features?

Besides that, one can also ask what the heuristic value of the notion "linguistic area" is (cf. Bisang 2006: 88–89 for a skeptical view vs. van Gijn & Wahlström 2023: 180 for a more optimistic one). If "linguistic areas" have any reality, linguists should be able to draw implications from statements like "the region *G* constitutes a linguistic area" or "language *L* belongs to the linguistic area *G*". In the ideal case, we expect languages belonging to a certain area to show similarities beyond the limited set of structural features used to define it, and, likewise, given that hardly any area is identified by an examination of all the languages and varieties populating a given region, we should expect yet unsampled varieties within or close to the geographical boundaries of the area to show at least some of the relevant structural features. These expectations, however, often turn out to be problematic, in particular when a putative area has been identified primarily on the basis of standard written varieties (as has been the case with the so-called "Standard Average European" and, at least partly, the Balkans). Taking into account non-standard varieties and the diastratic dimension in general can yield results significantly diverging from those based on standard languages, see e.g. Friedman (2008) on Balkan Slavic, Loporcaro (2011) on Romance, and Murelli & Kortmann (2011), Murelli 2011 on Europe in general.

3.2. Beyond "linguistic areas"

All the above-mentioned conceptual and methodological issues have been well-known and discussed for decades. Unsurprisingly, practitioners of areal typology have sought ways out of this impasse, both methodological and terminological. On the methodological side, the most important advancement has come from the integration of areal typology and contact linguistics. This involved a shift of perspective from linguistic areas as idealised categorically defined constructs to dynamic situations of language contact occurring in particular geographical regions, in particular historical periods, under particular sociocultural circumstances, and yielding particular structural outcomes, i.e. specific patterns of convergence between language varieties (for recent takes on this issue from a typological perspective see e.g. van Gijn 2020 and Di Garbo et al. 2021). Such contact situations can involve different languages at different periods, different types of bi- or multilingualism and their dynamics (e.g. in terms of language maintenance and language shift and their combinations, see e.g. Thomason & Kaufmann 1988 and much subsequent work), both horizontal (diatopic) and vertical (diastratic) dimensions (on these notions see Coseriu 1978: 223–225 and ***, this volume), as well as different "starting conditions" in terms of the degree of genealogical relatedness or structural similarity between languages prior to contact. It is therefore unsurprising that the outcomes of any history of contact reveals a "landscape" of both linguistic varieties and structural features looking anything but simple and straightforward. This has been shown to be true of the classic example of Sprachbund, i.e. the Balkans (see e.g. Friedman & Joseph 2017, 2025 and Joseph & Friedman, this volume, as well as Adamou & Sobolev (eds.) 2026), and certainly applies elsewhere.

Concomitantly, on the terminological side, various terms explicitly emphasising the language contact component of areality have been proposed. Those are "convergence area/zone", mentioned as a synonym to "linguistic area" already in Weinreich (1958: 379) and explicitly preferred to the latter in Bisang (2006: 75), and especially "contact superposition

zone", first introduced by Koptjevskaja-Tamm & Wälchli (2001: 624–626) with respect to the Circum-Baltic area, where not a single typologically outstanding structural trait can be argued to be present in all languages, but different subsets of languages show non-trivial similarities (see Seržant 2025 for a recent assessment). Koptjevskaja-Tamm & Wälchli (2001: 626) describe a contact superposition zone as involving "many layers of micro- and macro-contacts and mutual influences superimposed on each other over a long period of time", which "sometimes create an impression of an overall macro-contact among the languages in an area, which has not necessarily been there" (see e.g. Joseph & Friedman, this volume, on the Balkans). A largely synonymous term is "stratified convergence zone", see Drinka (this volume). My contention is that all so-called "linguistic areas" are in fact contact superposition zones that emerged as a result of an overlap of multilateral contact situations stratified on spatial, temporal and societal dimensions. Whether any such zone reveals a "neater" or "messier" distribution of any structural features is a largely accidental result of sociohistorical and geographical circumstances of its emergence and development.

Another relevant notion, which I have already mentioned in section 2 in the discussion of Johanna Nichols' approach, is that of "areal cline", i.e. "continuous increase and decrease of a continuous variable over a large area" (Wälchli 2005: 186), see Wälchli (2012) for a detailed discussion. Areal clines, particularly those defined over reasonably compact geographical regions, can be revealing of specific patterns of contact-induced diffusion. Importantly, even though typologists are probably more interested in clines involving many different and preferably unrelated languages, it is in fact most common to observe areal clines in so-called dialectal continua or groups of genealogically closely related varieties, where the spread of innovations through contact is facilitated by preexisting similarity (see e.g. contributions to Besters-Dilger et al. eds. 2014 as well as Jeszenszky et al. 2023). On closer inspection, however, such clines observed among closely related varieties can themselves turn out to be part of larger areal clines involving more distantly related or even unrelated languages. An interesting case in point is the distribution of accusative vs. genitive cases on the direct object of negated clauses in the Russian dialects; as shown by Seržant (2014) and Malysheva & Ronko (2025), this feature follows a clear north-to-south cline, with genitive being more frequent under negation in the northern Russian dialects. This dialectal cline is, at the same time, part of a much broader and more complex areal distribution of the genitive/partitive encoding of the direct object of negative sentences in parts of Eastern Europe (see Seržant 2015, Arkadiev & Kozhanov 2025 for details), which is most systematic in the Baltic Finnic languages that have served as adstratum or substratum precisely to the northern Russian dialects.

A prerequisite for identifying an areal cline is that the typological feature in question is defined in a gradual, rather than categorical, fashion. This relates not only or even not primarily to the nature of typological features themselves as to the types of data linguists take into consideration (cf. Wiemer & Arkadiev 2025: 21–23). Indeed, virtually any typological feature traditionally defined categorically can also be recast in such a way as to allow gradual assessment, sometimes in terms of the degree to which a certain feature is represented in the system (e.g. the distribution of prenasalisation across consonant phonemes, Donohue & Whiting 2011), but most notably in terms of frequency — either type frequency (e.g. of a certain phoneme in the lexicon, or of verbs with a particular valency frame among a sample of verbs) or token frequency of certain construction in a (parallel) corpus (see e.g. Wälchli 2009). A good example of such a gradual feature formingan areal cline is the encoding of experiencers in the European languages (Bossong 1998; Haspelmath 2001: 1495–1496), defined numerically as the proportion of predicates showing "inverted" encoding (i.e. similar to patient or recipient, rather than to the agent) among a fixed set of ten experiential predicates across 40 languages. Bossong has discovered a dispreference for inversion in the core Western European languages gradually decreasing towards the fringes of Europe (both westwards, e.g. Welsh and Icelandic,

and eastwards, e.g. Baltic and East Slavic), see Figure 4. Note that this feature is itself part of a broader "transitivity ratio" also showing a similar spatial distribution (Say 2014). As an example of a token-frequency based feature consider the occurrence of marked nominal plural in parallel texts studied by Wälchli (2009, 2012: 251–256) and showing a — certainly not exceptionless — west-to-east cline in Eurasia. Given the growing body of work in token-based typology (see in particular Levshina 2016, 2019; Levshina et al. 2023), we can expect more large- and small-scale areal clines to be revealed.

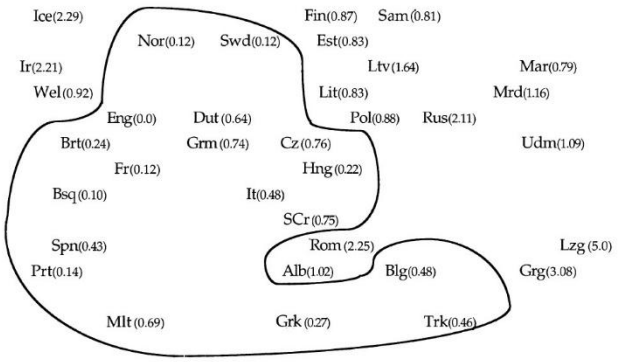

Figure 4. Inversion with experiencer predicates across Europe (Haspelmath 2001: 1496)

To summarise this discussion, I would like to suggest that the use of terms "area" and "areal" in linguistic contexts should return to its original understanding firmly grounded in geographical reality (e.g. Coseriu 1955). Recall the examples of "macroareal" and "microareal" distributions of selected typological features that were adduced at the beginning of section 2. When speaking about them, I only referred to the geographical patterns of languages with particular features, leaving aside the question about their possible "deeper" properties and origins (cf. Güldemann & Hammarström 2020: 59). My plea concerns exactly this: any non-random distribution of typological features observed in a particular geographical region is a valid object of inquiry as such, regardless of whether in the end this distribution turns out to be explainable by genealogical inheritance, contact-induced diffusion, universal tendencies or any interaction between them (see Wiemer et al. 2014, Wiemer & Arkadiev 2025 for a "triangulation" approach; for recent proposals of evaluating the relative contribution of these factors statistically, see e.g. Murawaki & Yamauchi 2018 and Ranacher et al. 2021). The ultimate goal of areal typology is to describe such distributions in an empirically adequate and comprehensive fashion and to arrive at unbiased explanations thereof, without a priori prioritising or excluding any possible factors. In the next section I'll briefly review some recent proposals to this end.

4. Some recent approaches to areal typology

Two very general kinds of approach in areal typology can be distinguished, which differ as to what they take as their starting point (see van Gijn &Wahlström 2023: 186–190 for a lucid discussion): non-linguistic factors, e.g. geography or cultural history, or linguistic factors, e.g. the distribution of typological features. Both types of approach have their advantages and limitations.

The most articulated geography-based approach, the "Predictive Areality Theory" (Bickel & Nichols 2006, Bickel 2017, 2020), has already been discussed in section 2. A proposal taking as a starting point the socio-cultural history of a certain region has been developed by Muysken (2008) with respect to the putative Atlantic linguistic area that emerged during the last millennium largely as a result of European colonisation of the Americas and of the African coast. In a recent article, van Gijn (2020) combines both socio-cultural and linguistic

perspectives and proposes a multi-layered "anatomy" of zones of convergence, including information about geography, social history, communication practices, and shared structural features and hypothesising that, at least in the ideal case, the extralinguistic factors should determine the structural ones. In section 3 I have already mentioned the potential circularity of the "classic" approach trying to establish Sprachbünde by means of analysing the distribution of selected structural properties in a group of selected languages. Several ways to circumvent such circularity have been proposed in the literature. Perhaps the most obvious of these follows the general contemporary trend of typology and shifts focus from comparison of languages to comparison of individual structural features themselves. Areal-typological studies of various concrete phenomena abound and some of them have already been mentioned above. Such studies can be further subclassified according to their geographical coverage (global, e.g. Nichols 1993 on ergativity, vs. local, e.g. Tuite 1999, also on ergativity, but in the languages of the Caucasus), their treatment of the phenomenon in question as "holistic" vs. "multivariate", i.e. consisting of potentially independent subfeatures (e.g. Arkadiev 2014 on prefixal perfectivisation or Matasović 2018 on agreement), as well as the definition of features as categorical vs. gradual, and employment of such data sources as (parallel) corpora (as e.g. Stolz et al. 2011 on total reduplication in the languages of Europe, or Wälchli 2024 on restrictives globally). One should also mention work on areal patterns in the lexicon (e.g. Koptjevskaja-Tamm & Liljegren 2017, Gast & Koptjevskaja-Tamm 2019) as well as in discourse (Beier at al. 2002, Dobrushina 2020), which constitutes an important complement to the general bias of (areal) typological studies towards phonology and grammar. An interesting approach considering both synchronic and historical data and revealing an increase of similarity through time is taken in Dedio et al. (2019) on the material of the languages of Britain and Ireland. This method, however, cannot be extended to most regions of the world, for which historical linguistic data is either scarce or altogether lacking.

Another feature-based approach avoids "cherry-picking" of "diagnostic" properties by applying statistical analysis to many different typological features, for instance, those from WALS. This approach has become prominent e.g. in the work by Mark Donohue and his associates (e.g. Donohue 2012, Donohue et al. 2008, 2011), who initially developed it in response to the claims by Dunn et al. (2005, 2008) that aggregate analysis of typological features can be revealing of phylogenetic relations between languages. As is convincingly shown e.g. in Donohue et al. (2011: 375–378), statistical methods of cluster analysis, such as NeighborNet, when applied to a set of WALS features coded for the Indo-European languages of Europe, produce results better corresponding to geographical proximity and contact between languages rather than to the phylogenetic trees built by the comparative method (e.g. Romanian is grouped with Albanian, and Bulgarian with Greek). This approach also allows researchers to test whether typological features from different domains pattern similarly across languages. Thus, Donohue (2012), again on the basis of the Indo-European languages of Europe, shows divergent behaviour of morphosyntactic and phonological features: while the former tend to cluster along genealogical lines with exceptions due to strong contact influence (as, again, with the languages of the Balkans), the latter appear to be sensitive to local contact events (as e.g. between Breton and French) as well as to "accidental similarities" (Donohue 2012: 108). Based on this, it is suggested that "typological feature analysis is best done not as a single large block of features, but broken down into many small subparts" (ibid.: 110) and that "different typological features yield different kinds of information, some more likely to reveal inheritance, and some more likely to indicate diffusion" (ibid.: 114), the point further developed in Donohue (2013).

The most impressive application of this method so far is reported in Kalyan & Donohue (2023) based on a database including more than 3000 languages and more than 300 morphosyntactic features, subjected to Factor Analysis of Mixed Data. The method revealed

that the most important (i.e. predictive) dimensions of cross-linguistic variation are verb-object order and elaboration of verbal morphology, both of which show strong geographical patterning, as well as a global east-west cline involving a number of features such as gender marking and ergativity. The method was able to detect several geographical clusters of languages occupying similar positions on certain dimensions, such as Western Old World (Europe, Arabia and North Africa), Mainland Southeast Asia, Pacific Northwest etc.

Other notable work taking a similar approach to areal typology includes, for instance, Cysouw (2011) on the relative typological exceptionality of geographically defined groups of languages (on the basis of WALS), Urban et al. 2019 on Middle and South America (based on 77 features across 44 languages), Ivani et al. 2020 on India (based on 217 variables and 27 languages) and Guzmán Naranjo & Mertner 2023 on Africa (645 languages, only features related to presence of certain phonemes). Although promising, such "big-data" approaches are dependent on the data annotation, correct choice of statistical methods and adequate interpretation of their results, which should be taken with caution pending replication studies (the point most recently emphasised by Becker & Guzmán Naranjo 2025).

One of the serious challenges for areal typological studies has been the integration of the information about geographical location of languages and linguistic communities as well as measuring distances between them, the latter often serving as proxies for the probability of language contact (on many problems of mapping languages in linguistic typology, see e.g. Mühlhäusler 2010; on the role of non-geographical factors such as writing or liturgical language in the diffusion of linguistic features, see e.g. Bisang 2013 and Drinka 2011; more generally on language mapping, see Rabanus 2020). Indeed, a glance at almost any linguistic map provided in works on (areal) typology reveals that every language is treated there as occupying a single point in space, in stark contrast to the tradition of dialectological maps. While probably not completely inadequate for regions with a dense population of languages spoken in a few nearby villages, this is certainly a rather gross idealisation with respect to many parts of the world where linguistic communities may occupy quite extended areas. What is even worse for areal typology, treating language locations as points does not allow researchers to model language spreads as well as situations where different linguistic communities occupy overlapping or intertwined geographical regions (and situations of multilingualism in general), which is particularly frequent in zones of intense convergence, many of which nowadays are localised in dense urban spaces (on this and other problematic aspects of traditional approaches to space-language relationships see e.g. Auer 2013 and Di Carlo 2022, 2024). On the other hand, it is obvious that trying to map many diverse languages in a geographically more adequate fashion is often unfeasible due to lack of reliable information as well as limitations of current statistical models and computing power.

Likewise, when assessing distances between points in space with exact coordinates representing idealised language locations, which distance metric should be chosen? Euclidian distances, even though simplest from the points of view of mathematics and computational implementation, in many cases only vaguely correspond to reality defined by the specifics of terrain, climate, types of roads and modes of transportation (see e.g. van Gijn et al. 2017 on the possible role of river systems in the linguistic landscape of South America). This matter has been seriously addressed only recently, see e.g. Guzmán Naranjo & Jäger (2024) and references therein.

5. Summary and outlook

In this by necessity incomplete and subjective overview, I hope to have presented the main trends and concepts of areal typology, focusing more on open questions and problematic issues than on results and achievements. As a linguist interested both in (large scale) typology and (small scale) areal and contact linguistics, I contend that these two domains of inquiry can and

should benefit from as much interaction and cross-fertilisation as possible. As was mentioned in the very beginning of this essay, almost any meaningful large-scale typological study inevitably has a geographical component, and the growing consensus among typologists is that this component is at least of interest and at most of crucial importance for understanding and explaining patterns of linguistic diversity. Likewise, many practitioners of more local areal linguistics have realised that setting the data of their languages of study against a wider geographical background, ultimately of continental and even global sizes, is, again, at least relevant and useful and at most crucially important for an adequate account and explanation. Similarly, integration of sociolinguistic, demographic, cultural, historical, anthropological and ecological information into studies dealing with distribution and diffusion of linguistic features is more often than not necessary for deeper understanding of causal factors and mechanisms explaining the patterns of linguistic diversity both on macro- and microlevels (see Di Grabo et al. 2025). This understanding can be also enhanced by recourse to more variegated types of data, such as (parallel) corpora, and more sophisticated methods designed to deal with non-categorical features. On the other hand, one should entertain necessary caution with respect to "high-level" statistical and computational methods, which, even if properly applied, yield results that are only as good as the input data and whose interpretation is only as good as the expertise of the linguist.

Acknowledgments
I thank the editors of the handbook for their generous invitation and patience, as well as Martin Haspelmath, Anastasia Panova, Sergey Say, Ilja Seržant, Björn Wiemer and three anonymous reviewers for their comments on the draft version of this article, and to Georgre Moroz for helping me to create the maps. None of the mentioned persons bears any responsibility for the views and interpretations expressed here.

References

Adamou, Evangelia & Andrey N. Sobolev (eds.). 2025. *Atlas of the Balkan Linguistic Area Online*. https://abla.cnrs.fr/

Adamou, Evangelia & Andrey N. Sobolev (eds.). 2026. *The Balkan Linguistic Area. Mapping Variation and Contact*. Berlin, Boston: De Gruyter Mouton.

Aikhenvald, Alexandra Yu. 2002. *Language Contact in Amazonia*. Oxford: Oxford University Press.

Arkadiev, Peter. 2014. Towards an areal typology of prefixal perfectivization. *Scando-Slavica* 60(2), 384–405. DOI: https://doi.org/10.1080/00806765.2014.984473

Arkadiev, Peter & Kirill Kozhanov. 2025. Object partitive of negation: An areal perspective. In Björn Wiemer, Peter Arkadiev, Rogier Blokland & Petar Kehayov (eds.), *Convergence and divergence in the eastern Circum-Baltic Area. A synthetic view, especially on Finnic, and case studies*, 513–569. Berlin, Boston: De Gruyter Mouton.

Auer, Peter. 2013. The geography of language: Steps toward a new approach. *Freiburger Arbeiten für Germanistische Linguistik* 16. https://www.sdd.uni-freiburg.de/fragl/2013.16

Bakker, Dik. 2010. Language sampling. In: Jae-Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*, 100–129. Oxford: Oxford University Press.

Becker, Laura & Matías Guzmán Naranjo. 2025. Replication and methodological robustness in quantitative typology. *Linguistic Typology* 29(3), 463–505.

Beier, Christine, Lev Michael & Joel Scherzer. 2002. Discourse forms and processes in Lowland South America: An areal-typological perspective. *Annual Review of Anthropology* 31, 121–145.

Besters-Dilger, Juliane, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.). 2014. *Congruence in Contact-induced Language Change (Language Families, Typological Resemblance, and Perceived Similarity)*. Berlin, Boston: De Gruyter.

Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11, 239–251.

Bickel, Balthasar. 2013. Distributional biases in language families. In: Balthasar Bickel, Lenore Grenoble, David Peterson & Alan Timberlake (eds.), *Language Typology and Historical Contingency: In Honor of Johanna Nichols*, 415–444. Amsterdam, Philadelphia: John Benjamins.

Bickel, Balthasar. 2017. Areas and universals. In: Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 40–54. Cambridge: Cambridge University Press.

Bickel, Balthasar. 2020. Large and ancient linguistic areas. In: Mily Crevels & Pieter Muysken (eds.), *Language Dispersal, Diversification, and Contact. A Global Perspective*, 78–100. Oxford: Oxford University Press.

Bickel, Balthasar & Johanna Nichols. 2002. Autotypologizing databases and their use in fieldwork. In: Peter Austin, Helen Dry and Peter Witternburg (eds.), *Proceedings of the International LREC Workshop on Resources and Tools in Field Linguistics, Las Palmas, 26–27 May 2002*.

Bickel, Balthasar & Johanna Nichols. 2006. Oceania, the Pacific Rim and the theory of linguistic areas. In: *Proceedings of the 32nd Annual Meeting of the Berkeley Linguistics Society*, https://doi.org/10.3765/bls.v32i2.3488

Bisang, Walter. 2006. Linguistic areas, language contact and typology: Some implications from the case of Ethiopia as a linguistic area. In: Yaron Matras, April McMahon & Nigel Vincent (eds.), *Linguistic Areas: Convergence in Historical and Typological Perspective*, 75–98. Hampshire: Palgrave MacMillan.

Bisang, Walter. 2010. Areal language typology. In: Peter Auer & Jürgen E. Schmidt (eds.). *Language and Space. An International Handbook of Linguistic Variation*, Vol. 1: *Theories and Methods*, 419–440. Berlin, New York: Mouton de Gruyter.

Bisang, Walter. 2013. Language contact between geographic and mental space. In: Peter Auer, Martin Hilpert, Anja Stukenbrock & Benedikt Szmrecsanyi (eds.), *Space in Language and Linguistics. Geographical, Interactional, and Cognitive Perspectives*, 61–100. Berlin, Boston: Walter de Gruyter.

Boas, Franz. 1917. Introductory. *International Journal of American Linguistics* 1(1), 1–8.

Boas, Franz. 1937. The diffusion of cultural traits. *Social Research: An International Quarterly of Political and Social Science* 4(3), 286–259.

Boas, Franz. 1929. Classification of American Indian languages. *Language* 5(1), 1–7.

Bossong, Georg. 1998. Le marquage de l'expérient dans les langues d'Europe. In: Jack Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, 259-294. Berlin, New York: Mouton De Gruyter.

Campbell, Lyle. 2002. What drives linguistic diversification and language spread? In: Peter Bellwood & Colin Renfrew (eds.), *Examining the Farming/Language Dispersal Hypothesis*, 49–63. Cambridge: McDonald Institute for Archaeological Research.

Campbell, Lyle. 2006. Areal linguistics: A closer scrutiny. In: Yaron Matras, April McMahon & Nigel Vincent (eds.), *Linguistic Areas: Convergence in Historical and Typological Perspective*, 1–31. Hampshire: Palgrave MacMillan.

Campbell, Lyle. 2017. Why is it so hard to define a linguistic area? In: Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 19–39. Cambridge: Cambridge University Press.

Campbell, Lyle, Terrence Kaufman & Thomas C. Smith-Stark. 1986. Mesoamerica as a linguistic area. *Language* 62(3), 530–570.

Chirikba, Viacheslav. 2008. The problem of the Caucasian Sprachbund. In: Pieter Muysken (ed.), *From Linguistic Areas to Areal Linguistics*, 25–93. Amsterdam, Philadelphia: John Benjamins.

Comrie, Bernard. 2020. Mapping the world's languages: From data via purpose to representation. In: Stanley D. Brunn & Roland Kehrein (eds.), *Handbook of the Changing World Language Map*, 61–76. Cham: Springer.

Coseriu, Eugenio. 1955. La geografía lingüistica. *Revista de la Facultad de Humanidades y Ciencia*, 14, 29–69. Montevideo: Universidad de la República.

Coseriu, Eugenio. 1978. Einführung in die strukturelle Betrachtung des Wortschatzes. In: Horst Geckeler (Hrsg.), *Strukturelle Bedeutungslehre*, 193–238. Darmstadt: Wissenschaftliche Buchgesellschaft.

Cysouw, Michael. 2011. Quantitative explorations of the world-wide distribution of rare characteristics, or: the exceptionality of northwestern European languages. In: Horst Simon and Heike Wiese (eds.), *Expecting the Unexpected: Exceptions in Grammar*, 411–431. Berlin: De Gruyter Mouton.

Dahl, Östen. 2001. Principles of areal typology. In: Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals*, 1456–1470. Berlin: Walter de Gruyter.

Dahl, Östen (ed.) 2000. *Tense and Aspect in the Languages of Europe*. Berlin, New York: Mouton de Gruyter.

Danylenko, Andrii. 2019. Balkanisms and Carpathianisms or, Carpathian Balkanisms? In: Andrii Danylenko & Motoki Nomachi (eds.), *Slavic on the Language Map of Europe: Historical and Areal-Typological Dimensions*, 347–383. Berlin, Boston: De Gruyter.

Dedio, Stefan, Peter Ranacher & Paul Widmer. 2019. Evidence for Britain and Ireland as a linguistic area. *Language* 95(3), 498–522.

Di Carlo, Pierpaolo. 2022. The georgraphic sides of small-scale multilingualism: New challenges in linguistic cartography. In: Greg Nieddt (ed.), *New Directions in Linguistic Geography. Exploring Articulations of Space*, 49–85. Singapore: Palgrave Macmillan.

Di Carlo, Pierpaolo. 2024. Language maps and sociolinguistic data — Developing linguistic cartography of Bantoid languages. *Afrika und Übersee* 97, 87–113.

Di Garbo, Francesca, Eri Kashima, Ricardo Napoleão de Souza & Kaius Sinnemäki. 2021. Concepts and methods for integrating language typology and sociolinguistics. In: Silvia Ballarè & Guglielmo Inglese (eds.), *Tipologia e Sociolinguistica: Verso un approccio integrato allo studio della variazione: Atti del Workshop della Società Linguistica Italiana 20 settembre 2020*, 143–176. Milano: Officinaventuno.

Di Garbo, Francesca, Kaius Sinnemäki & Eri Kashima. 2025. Towards greater social anchoring in language typology. *Linguistic Typology at the Crossroads* 5(2), 1–35.

Dobrushina, Nina. 2020. Optative in the Caucasus. Evidence for a linguistic or discourse area? Presentation at the 53rd Annual Meeting of the Societas Linguistica Europaea, online, 26 August – 1 September 2020.

Donohue, Mark. 2012. Typology and areality. *Language Dynamics and Change* 2, 98–116.

Donohue, Mark. 2013. Who inherits what, when? Towards a theory of contact, substrates and superimposition zones. In: Balthasar Bickel, Lenore Grenoble, David Peterson & Alan Timberlake (eds.), *Language Typology and Historical Contingency. In Honor of Johanna Nichols*, 219–239. Amsterdam, Philadelphia: John Benjamins.

Donohue, Mark, Simon Musgrave, Bronwen Whiting & Søren Wichmann. 2011. Typological feature analysis models linguistic geography. *Language* 87(2), 369–383.

Donohue, Mark & Bronwen Whiting. 2011. Quantifying areality: A study of prenasalization in Southeast Asia and New Guinea. *Linguistic Typology* 15, 101–121.

Donohue, Mark, Søren Wichmann & Mihai Albu. 2011. Typology, areality, and diffusion. *Oceanic Linguistics* 47(1), 223–232.

Drinka, Bridget. 2011. The sacral stamp of Greek: Periphrastic constructions in New Testament translations of Latin, Gothic, and Old Church Slavonic. *Oslo Studies in Language* 3(3), 41–73.

Drinka, Bridget. 2017. *Language Contact in Europe. The Periphrastic Perfect through History*. Cambridge: Cambridge University Press.

Dryer, Matthew S. 1989. Large linguistic areas and language sampling. *Studies in Language* 13(2), 257–292.

Dryer, Matthew S. 2013. Order of adjective and noun. In: Matthew S. Dryer & Martin Haspelmath (eds.) *WALS Online*. Available online at http://wals.info/chapter/87, accessed on 2025-01-13.

Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *WALS Online*. Available online at https://wals.info/

Dunn, Michael, Angela Terrill, Ger Reesink, Robert A. Foley & Stephen C. Levinson. 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science* 309: 2072–2075.

Dunn, Michael, Stephen C. Levinson, Eva Lindström, Ger Reesink & Angela Terrill. 2008. Structural phylogeny in historical linguistics: methodological explorations applied in Island Melanesia. *Language* 84(4), 710–759.

Emeneau, Murray B. 1956. India as a linguistic area. *Language* 32(1), 3–16.

Enfield, Nick J. 2005. Areal linguistics and Mainland Southeast Asia. *Annual Review of Anthropology* 34, 181–206.

Fleischer, Jürg. 2004. A typology of relative clauses in German dialects. In: Bernd Kortmann (ed.), *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*, 211–243. Berlin, New York: Mouton de Gruyter.

Friedman, Victor. 2008. Balkan Slavic dialectology and Balkan linguistics: Periphery as center. In: Chr. Y. Bethin (ed.), *American Contributions to the 14th International Congress of Slavists, Ohrid, September 2008*, 131–148. Bloomington, IN: Slavica.

Friedman, Victor & Brian D. Joseph. 2017. Reassessing sprachbunds: A view from the Balkans. In: Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 55–87. Cambridge: Cambridge University Press.

Friedman, Victor A. & Brian D. Joseph. 2025. *The Balkan Languages*. Cambridge: Cambridge University Press.

Gast, Volker & Maria Koptjevskaja-Tamm. 2019. The areal factor in lexical typology. In: Daniël van Olmen, Tanja Mortelmans & Frank Brisand (ed.) *Aspects of Linguistic Variation. Studies in Honor of Johan van der Auwera*, 43–81. Berlin, Boston: De Gruyter Mouton.

Grzega, Joachim (ed.) 2025. *The Routledge Handbook of Eurolinguistics*. London, New York: Routledge.

Güldemann, Tom. 2008. The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa. In: Bernd Heine & Derek Nurse (eds.), *A Linguistic Geography of Africa*, 151–185. Cambridge: Cambridge University Press.

Güldemann, Tom. 2018. Areal linguistics beyond contact, and linguistic areas in Afrabia. In: Tom Güldemann (ed.), *The Languages and Linguistics of Africa*, 448–545. Berlin, Boston: De Gruyter Mouton.

Güldemann, Tom & Harald Hammarström. 2020. Geographical axis effects in large-scale linguistic distributions. In: Mily Crevels & Pieter Muysken (eds.), *Language Dispersal, Diversification, and Contact. A Global Perspective*, 58–77. Oxford: Oxford University Press.

Guzmán Naranjo, Matías & Laura Becker. 2022. Statistical bias control in typology. *Linguistic Typology* 26(3), 605–670.

Guzmán Naranjo, Matías & Gerhard Jäger. 2024. Euclide, the crow, the wolf and the pedestrian: distance metrics for linguistic typology. *Open Research Europe*, 3:104. https://doi.org/10.12688/openreseurope.16141.2

Guzmán Naranjo, Matías & Miri Mertner. 2023. Estimating areal effects in typology: a case study of African phoneme inventories. *Linguistic Typology* 27(2), 455–480.

Hammarström, Harald & Mark Donohue. 2014. Some principles of the use of macro-areas in typological comparison. *Language Dynamics and Chang*e 4, 167–187.

Hartmann, Frederik & Johanna Nichols. 2025. Geospatial effects on phonological complexity in the world's languages. *Linguistic Typology* 29(3), 627–670.

Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. In: Haspelmath, Martin, Ekkehard König, Wulf Österreicher & Wolfgang Raible (eds.). *Language Typology and Language Universals. An International Handbook*, 1492–1510. Berlin, New York: Mouton de Gruyter.

Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds.). 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Heath, Jeffrey. 1978. *Linguistic Diffusion in Arnhem Land*. Canberra: Australian Institute of Aboriginal Studies.

Heggarty, Paul. 2017. Towards a (pre)history of linguistic convergence areas: correlates of genetics, archaeology, history and geography. In: *Mémoires de la Société de linguistique de Paris*, *T. 24. Diffusion: Implantation, affinités, convergence*, 135–178. Louvain: Peeters.

Heine, Bernd & Tania Kuteva. 2006. *The Changing Languages of Europe*. Oxford: Oxford University Press.

Hickey, Raymond. ed. 2017. *The Cambridge Handbook of Areal Linguistics*. Cambridge: Cambridge University Press.

Huisman, John L.A., Asifa Majid & Roeland van Hout. 2019. The geographical configuration of a language area influences linguistic diversity. *PLos ONE* 14(6):e0217363. https://doi.org/10.1371/journal.pone.0217363

Inman, David., Natalia Chousou-Polydouri, Marine Vuillermet et al. 2025. The Areal Typology of Languages of the Americas (ATLAs) database. *Scientific Data* 12, 933. https://doi.org/10.1038/s41597-025-05169-4Ivani, Jessika K., Netra P. Paudyal & John Peterson. 2020. Indo-Aryan — a house divided? Evidence for the east-west Indo-Aryan divide and its significance for the study of northern South Asia. *Journal of South Asian Languages and Linguistics* 7(2), 287–326.

Jakobson, Roman. 1931. Über die phonologischen Sprachbünde. *Travaux du cercle linguistique de Prague* 4, 234–240.

Jakobson, Roman. 1972[1938]. On the theory of phonological associations among languages. In: Allan R. Keiler (ed.), *A Reader in Historical and Comparative Linguistics*, 241–252. New York etc.: Holt, Rinehart & Winston. (1st published in *Proceedings of the Fourth International Congress of Linguists*, 1938, 45–48).

Jeszenszky, Péter, Anja Hasse & Philipp Stöckle. 2023. Dialect areas and contact dialectology. In: Rik van Gijn, Hanna Ruch, Max Wahlström & Anja Hasse (eds.), *Language Contact: Bridging the Gap between Individual Interactions and Areal Patterns*, 135–177. Berlin: Language Science Press.

Kalyan, Siva & Mark Donohue. 2023. The dimensions of morphosyntactic variation: Whorf, Greenberg and Nichols were right. *Linguistic Typology at the Crossroads* 3(2), 132–190.

Koptjevskaja-Tamm, Maria & Bernhard Wälchli. 2001. The Circum-Baltic languages: An areal-typological approach. In: Östen Dahl & Maria Koptjevskaja-Tamm (eds.). *The*

Circum-Baltic Languages. Vol. 2. *Grammar and Typology*, 615–750. Amsterdam, Philadelphia: John Benjamins.

Koptjevskaja-Tamm, Maria & Henrik Liljegren. 2017. Semantic patterns from an areal perspective. In: Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 205–236. Cambridge: Cambridge University Press.

Levshina, Natalia. 2016. Why we need a token-based typology: A case study of analytic and lexical causatives in fifteen European languages. *Folia Linguistica* 50(2), 507–542.

Levshina, Natalia. 2019. Token-based typology and word order entropy: A study based on Universal Dependencies. *Linguistic Typology* 23(3), 533–572.

Levshina, Natalia, Savithry Namboodiripad et al. 2023. Why we need a gradient approach to word order. *Linguistics* 61(4), 825–883.

Liljegren, Henrik. 2020. The Hindu Kush–Karakorum and linguistic areality. *Journal of South Asian Languages and Linguistics* 7(2), 239–285.

Loporcaro, Michele. 2011. Two euroversals in a global perspective: auxiliation and alignment. In: Peter Siemund (ed.), *Linguistic Universals and Language Variation*, 57–91. Berlin, New York: Mouton de Gruyter.

Maddieson, Ian. 2013. Presence of uncommon consonants. In: Matthew S. Dryer & Martin Haspelmath (eds.) *WALS Online*. Available online at http://wals.info/chapter/19, accessed on 2025-01-13.

Malysheva, Anna & Roman Ronko. 2025. Case selection for the direct object of negated verbs in Russian dialects. To appear in Björn Wiemer, Peter Arkadiev, Rogier Blokland & Petar Kehayov (eds.), *Convergence and divergence in the eastern Circum-Baltic Area. A synthetic view, especially on Finnic, and case studies*. Berlin, Boston: De Gruyter Mouton.

Masica, Colin. 1976. *Defining a Linguistic Area: South Asia*. Chicago, London: The University of Chicago Press.

Masica, Colin. 2001. The definition and significance of linguistic areas: Methods, pitfalls and possibilities (with special reference to the validity of South Asia as a linguistic area). In: P. Bhaskararao & K.V. Subbarao (eds.), *The Yearbook of South Asian Languages and Linguistics. 2001. Tokyo Symposium on South Asian Languages. Contact, Convergence and Typology*, 205–267. New Delhi etc.: Sage Publications.

Matasović, Ranko. 2018. *An Areal Typology of Agreement Systems*. Cambridge: Cambridge University Press.

McConvell, Patrick. 1996. Backtracking to Babel: the chronology of Pama-Nyungan expansion in Australia. *Archaeology in Oceania* 31, 125–144.

Miestamo, Matti. 2005. *Standard Negation. The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Berlin, New York: Mouton de Gruyter.

Miestamo, Matti, Dik Bakker & Antti Arppe. 2016. Sampling for variety. *Linguistic Typology* 20(2), 233–296.

Mešeska Tomić, Olga. 2008. An integrated areal-typological approach. Local convergence of morphosyntactic features in the Balkan Sprachbund. In: Pieter Muysken (ed.), *From Linguistic Areas to Areal Linguistics*, 181–220. Amsterdam, Philadelphia: John Benjamins.

Moroz, George. 2017. lingtypology: easy mapping for Linguistic Typology. https://CRAN.R-project.org/package=lingtypology .

Moroz, George, Michael Daniel, Konstantin Filatov, Timur Maisak, Timofey Mukhin, Irina Politova, Elena Shvedova, Samira Verhees & Chiara Naccarato. 2026. *Typological Atlas of the Languages of Daghestan (TALD)*, v. 2.0.1. Moscow: Linguistic Convergence Laboratory, HSE University. DOI: 10.5281/zenodo.6807070. https://lingconlab.ru/tald

Mühlhäusler, Peter. 2010. Mapping linguistic typology. In: Alfred Lameli, Roland Kehrein & Stefan Rabanus (eds.) *Language and Space. An International Handbook of Linguistic*

Variation. Vol. 2. Language Mapping. Part I, 355–374. Berlin, New York: De Gruyter Mouton.

Murawaki, Yugo & Kenji Yamauchi. 2018. A statistical model for the joint inference of vertical stability and horizontal diffusibility of typological features. *Journal of Language Evolution*, 13–25.

Murelli, Adriano. 2011. *Relative Constructions in European Non-Standard Varieties*. Berlin, Boston: De Gruyter Mouton.

Murelli, Adriano & Bernd Kortmann. 2011. Non-standard varieties in the areal typology of Europe. In: Bernd Kortmann & Johan van der Auwera (eds.), *The Languages and Linguistics of Europe: A Comprehensive Guide*, 525–544. Berlin, Boston: De Gruyter Mouton.

Muysken, Pieter. 2008. Introduction. Conceptual and methodological issues in areal linguistics. In: Pieter Muysken (ed.), *From Linguistic Areas to Areal Linguistics*, 1–23. Amsterdam, Philadelphia: John Benjamins.

Nekula, Marek. 2003. Sprachbund und Sprachtyp. In: Marek Nekula (Hrsg.), *Prager Strukturalismus: Methodologische Grundlagen*, 79–103. Heidelberg: Winter.

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago, London: The University of Chicago Press.

Nichols, Johanna. 1993. Ergativity and linguistic geography. *Australian Journal of Linguistics* 13(1), 39–89.

Nichols, Johanna. 1997. Modeling ancient population structure and movement in linguistics. *Annual Review of Anthropology* 26, 359–384.

Nichols, Johanna. 2013. The vertical archipelago: Adding the third dimension to linguistic geography. In: Peter Auer, Martin Hilpert, Anja Stukenbrock & Benedikt Szmrecsanyi (eds.), *Space in Language and Linguistics. Geographical, Interactional, and Cognitive Perspectives*, 38–60. Berlin, Boston: Walter de Gruyter.

Nichols, Johanna. 2015. Types of spread zones. Open and closed, horizontal and vertical. In: Rik De Busser & Randy J. LaPolla (eds.), *Language Structure and Environment. Social, Cultural, and Natural Factors*, 261–286. Amsterdam, Philadelphia: John Benjamins.

Nichols, Johanna. 2017. Person as an inflectional category. *Linguistic Typology* 21(3), 387–456.

Nichols, Johanna. 2020. Dispersal patterns shape areal typology. In: Mily Crevels & Pieter Muysken (eds.). *Language Dispersal, Diversification, and Contact*, 25–43. Oxford: Oxford University Press.

Nichols, Johanne & Balthasar Bickel. 2013. Possessive classification. In: *WALS Online*. Available online at http://wals.info/chapter/59, accessed on 2025-01-23.

Nichols, Johanna, David A. Peterson & Jonathan Barnes. 2004. Transitivizing and detransitivizing languages. *Linguistic Typology* 8(2), 149–211.

Plank, Frans (ed.) 2003. *Noun Phrase Structure in the Languages of Europe*. Berlin, New York: Mouton de Gruyter.

Rabanus, Stefan. 2020. Language mapping worldwide: Methods and traditions. In: Stanley D. Brunn & Roland Kehrein (eds.), *Handbook of the Changing World Language Map*, 103–129. Cham: Springer.

Ranacher, Peter, Nico Neureiter, Rik van Gijn, Barbara Sonnenhauser, Anastasia Escher, Robert Weibel, Pieter Muysken & Balthasar Bickel. 2021. Contact-tracing in cultural evolution: a Bayesian mixture model to detect geographic areas of language contact. *Journal of the Royal Society Interface* 18: 20201031.

Rijkhoff, Jan, Dik Bakker, Kees Hengeveld & Peter Kahrel. A method of language sampling. *Studies in Language* 17(1), 169–203.

Sandfeld, Kristian. 1930. *Linguistique balkanique. Problèmes et résultats*. Paris: Klincksieck.

Say, Sergey. 2014. Bivalent verb classes in the languages of Europe: A quantitative typological study. *Language Dynamics and Change* 4, 116–166.

Seržant, Ilja. 2014. The independent partitive genitive in North Russian. In: Ilja A. Seržant & Björn Wiemer (eds.), *Contemporary Approaches to Dialectology: The area of North, Northwest Russian and Belarusian vernaculars*, 270–329. Bergen: John Grieg AS.

Seržant, Ilja A. 2015. The independent partitive as an Eastern Circum-Baltic isogloss. *Journal of Language Contact* 8. 341–418.

Seržant, Ilja. 2025. Circum-Baltic convergence area. In: In Marc L. Greenberg (ed.), *Encyclopedia of Slavic Languages and Linguistics* Online. DOI: https://doi.org/10.1163/2589-6229_ESLO_COM_048365

Shagal, Ksenia, Max Wahlström, Johanna Nichols & Stefan Dedio (under revision). (Non-)finiteness in clause combining in northern Eurasia. *Linguistic Typology at the Crossroads*.

Stassen, Leon. 2010. The problem of cross-linguistic identification. In: Jae-Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 90–99.

Stolz, Thomas. 2002. No Sprachbund beyond this line: On the age-old discussion of how to define a linguistic area. In: Paolo Ramat & Thomas Stolz (eds.), *Mediterranean languages: Papers from the MEDTYP workshop, Tirrenia, June 2000*, 259–281. Bochum: Brockmeyer.

Stolz, Thomas, Sonja Kettler, Cornelia Stroh & Aina Urdze. 2008. *Split Possession. An Areal-Linguistic Study of the Alienability Correlation and Related Phenomena in the Languages of Europe*. Amsterdam, Philadelphia: John Benjamins.

Stolz, Thomas, Nataliya Levkovych, Aina Urdze, Julia Nintemann & Maja Robbers. 2017. *Spatial Interrogatives in Europe and Beyond: Where, Whither, Whence*. Berlin, Boston: De Gruyter Mouton.

Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2006. *On Comitatives and Related Categories: A Typological Study with Special Reference to the Languages of Europe*. Berlin, New York: Mouton de Gruyter.

Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2011. *Total Reduplication. The Areal Linguistics of a Potential Universal*. Berlin: Akademie-Verlag.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, Los Angeles, London: University of California Press.

Toman, Jindřich. 1995. *The Magic of a Common Language: Jakobson, Mathesius, Trubetzkoy, and the Prague Linguistic Circle*. Cambridge, MA: MIT Press.

Tosco, Mauro. 2000. Is there an "Ethiopian" language area? *Anthropological Linguistics* 42(3), 329–365.

Trubetzkoy, Nikolai S. 1923. Vavilonskaja bašnja i smešenie jazykov [The Tower of Babel and the mixing of languages]. *Evrazijskij vremennik* 3, 107–124.

Trubetzkoy, Nikolai S. 1930. Proposition 16. Über den Sprachbund. In: *Actes du premier congrès international des linguistes à la Haye, du 10-15 avril 1928*, 17–18. Leiden: A. W. Sijthoff.

Tuite, Kevin. 1999. The myth of the Caucasian Sprachbund: The case of ergativity. *Lingua* 108, 1–26.

Urban, Matthias. 2019. Is there a Central Andean linguistic area? A view from the perspective of the "minor" languages. *Journal of Language Contact* 12, 217–304.

Urban, Matthias. 2020. Mountain linguistics. *Language and Linguistics Compass* 14:e12393. https://doi.org/10.1111/lnc3.12393

Urban, Matthias. 2021. The geography and development of language isolates. *Royal Society Open Science* 8:202232. https://doi.org/10.1098/rsos.202232

Urban, Matthias, Hugo Reyes-Centeno, Kate Bellamy & Matthias Pache. 2019. The areal typology of western Middle and South America: Towards a comprehensive view. *Linguistics* 57(6), 1403–1463.

van der Auwera, Johan. 1998. Revisiting the Balkan and Meso-American linguistic areas. *Language Sciences* 20(3), 259–270.
van der Auwera, Johan (ed.) 1998. *Adverbial Constructions in the Languages of Europe*. Berlin, New York: Mouton de Gruyter.
van Gijn, Rik. 2020. Separating layers of information: The anatomy of contact zones. In: Norval Smith, Tonjes Veenstra & Enoch O. Aboh (eds.), *Advances in contact linguistics: In honour of Pieter Muysken*, 162–178. Amsterdam, Philadelphia: John Benjamins.
van Gijn, Rik, Harald Hammarström, Simon van de Kerke, Olga Krasnoukhova & Pieter Muysken. 2017. Linguistic areas, linguistic convergence and river systems in South America. In: Raymond Hickey (ed.), *The Cambridge Handbook of Areal Linguistics*, 964–996. Cambridge: Cambridge University Press.
van Gijn, Rik & Max Wahlström. 2023. Linguistic areas. In: Rik van Gijn, Hanna Ruch, Max Wahlström & Anja Hasse (eds.), *Language Contact. Bridging the Gap between Individual Interactions and Areal Patterns*, 179–219. Berlin: Language Science Press.
Van Pottelberge, Jeroen. 2001. Sprachbünde: Beschreiben sie Sprachen oder Linguisten? *Linguistik Online* 8(1). https://doi.org/10.13092/lo.8.978
Velten, Harry V. 1943. The Nez Perce verb. *The Pacific Northwest Quarterly* 34(3), 271–292.
Velupillai, Viveka. 2012. *An Introduction to Linguistic Typology*. Amsterdam, Philadelphia: John Benjamins.
Wälchli, Bernhard. 2005. *Co-Compounds and Natural Coordination*. Oxford: Oxford University Press.
Wälchli, Bernhard. 2009. Data reduction typology and the bimodal distribution bias. *Linguistic Typology* 13(1), 77–94.
Wälchli, Bernhard. 2012. Grammaticalization clines in space: zooming in to synchronic traces of diffusion processes. In: Björn Wiemer, Bernhard Wälchli & Björn Hansen (eds.), *Grammatical Replication and Borrowability in Language Contact*, 233–273. Berlin, Boston: Mouton de Gruyter.
Wälchli, Bernhard. 2024. We need world-wide corpus-based typology: A parallel corpus study of restrictives ('only'). *Travaux neuchâtelois de linguistique* 79, 69–157.
Weinreich, Uriel. 1958. On the compatibility of genetic relationship and convergent development. *Word* 14(2-3), 374–379.
Wiemer, Björn. 2019. *Matrëška* and areal clusters involving varieties of Slavic. On methodology and data treatment. In: Andrii Danylenko & Motoki Nomachi (eds.), *Slavic on the Language Map of Europe*, 21–61. Berlin, Boston: De Gruyter Mouton.
Wiemer, Björn & Peter Arkadiev. 2025. Introduction. In: Björn Wiemer, Peter Arkadiev, Rogier Blokland & Petar Kehayov (eds.), *Convergence and divergence in the eastern Circum-Baltic Area. A synthetic view, especially on Finnic, and case studies*, 1–67. Berlin, Boston: De Gruyter Mouton.
Wiemer, Björn, Ilja Seržant & Aksana Erker. 2014. Convergence in the Baltic-Slavic contact zone (Triangulation approach). In: Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder & Achim Rabus (eds.), *Congruence in Contact-induced Language Change (Language Families, Typological Resemblance, and Perceived Similarity)*, 15–42. Berlin, Boston: De Gruyter.
Witzlack-Makarevich and Balthasar Bickel (eds.). 2019. *Argument Selectors. A New Perspective on Grammatical Relations*. Amsterdam, Philadelphia: John Benjamins.