ScienceAdvances AAAS

Submitted manuscript

Societies of strangers do not speak grammatically simpler languages

Authors

Olena Shcherbakova,[1]* Susanne Maria Michaelis,[1] Hannah J. Haynie,[2] Sam Passmore,[3] Volker Gast,[4] Russell D. Gray,[1,5] Simon J. Greenhill,[6,1] Damián E. Blasi,[7, 1,8] & Hedvig Skirgård[1]

Affiliations

[1]Department of Linguistic and Cultural Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany.

[2]Department of Linguistics, University of Colorado Boulder, Boulder, CO 80309.

[3]Evolution of Cultural Diversity Initiative, School of Culture, History and Language, College of Asia and the Pacific, The Australian National University, Canberra, ACT 2601, Australia.

[4]Department of English and American Studies, Friedrich-Schiller University of Jena, Jena 07745, Germany.

[5]School of Psychology, University of Auckland, 1010 Auckland, New Zealand.

[6]School of Biological Sciences, University of Auckland, 1010 Auckland, New Zealand.

[7]Department of Human Evolutionary Biology, Peabody Museum, Harvard University, Cambridge, MA 02138.

[8]Human Relations Area Files, Yale University, New Haven, CT 06520.

*shcherbakova@shh.mpg.de

Abstract

Many recent proposals claim that languages adapt to their environments. The Linguistic Niche hypothesis claims that languages with numerous native speakers and substantial proportions of non-native speakers (societies of strangers) will tend to lose grammatical distinctions. In contrast, languages in small, isolated communities should maintain or expand their range of grammatical markers. Here, we test such claims using a new global dataset of grammatical structures - Grambank. We model the impact of the number of native speakers, the proportion of non-native speakers, the number of linguistic neighbors, and the status of a language on grammatical complexity while controlling for spatial and phylogenetic autocorrelation. We deconstruct "grammatical complexity" into two separate dimensions: (i) how much morphology a language has ("fusion"), and (ii) the amount of information obligatorily encoded in the grammar ("informativity"). We find several instances of weak positive associations but no inverse correlations between grammatical complexity and sociodemographic factors. Our findings cast doubt on the widespread assumption that grammatical complexity is shaped by the sociolinguistic environment.

Teaser

This study uses a dataset of 1,314 languages to test the claim that grammatical complexity is reduced in societies of strangers.

Introduction

Societies vary greatly in their size, homogeneity, and the degree of contact with other societies. The variation in these properties is captured in the continuum between two extremes: "societies of intimates" (esoteric societies) and "societies of strangers" (exoteric societies) (*1*–*4*). Societies of intimates are small-sized tight-knit homogenous groups where members share high amounts of knowledge about community life and do not engage much with outsiders (*1*, *2*, *5*–*8*). On another extreme, we encounter societies of strangers: large heterogeneous groups with substantial proportions of outsiders (either people using a different language altogether or at least outsiders to the local community), loose networks, and, as a result, lower amounts of shared communal information. We refer to this continuum that spans from societies of intimates to societies of strangers as exotericity. Within this continuum, societies low in exotericity are prototypical societies of intimates (esoteric societies), and those high in exotericity correspond to the characteristics of societies of strangers.

Different degrees of exotericity in societies have been hypothesized to shape the communication between individuals – ultimately resulting in observable effects on grammatical structures. Two prominent pathways link exotericity and language structure. Firstly, the members of homogenous esoteric societies rarely communicate with outsiders, and hence, the languages in such societies are acquired and used almost exclusively by members of these societies. This lack of contact with non-native speakers has been claimed to shape languages such that they develop and retain more obligatory explicit grammatical marking (*1*, *5*, *9*). For example, Tariana, an endangered Arawakan language in the Amazonas has an evidential marking system: the verbs carry grammatical information that distinguishes between situations where the speaker has seen or heard the action they report, or where this action is inferred or assumed from second- or third-hand information (*10*). This grammatical feature has been suggested to occur in low-exotericity languages rather than highly exoteric ones (*11*).

Secondly, the social setting of exoteric communities with high proportions of outsiders and degrees of contact with non-native L2 speakers has been proposed to drive morphological simplification in languages (*5*, *6*). L2 speakers find it especially difficult to process and produce phonologically fused grammatical structures, such as case endings and verbal agreement markers (*5*, *6*, *12*, *11*, *13*). Hence, such languages have been suggested to undergo a process of simplification, such as loss of morphological categories and agreement. For instance, since the Old English period, English has lost the adjective agreement in case, number, and gender as well as the nominal case distinctions, which has been linked to the adoption of English by non-native speakers (*11*). Similarly, it has been proposed that gender systems, another feature more typical of low-exotericity languages (*11*), tend to reduce in languages that undergo contact with other languages, especially those without gender, i.e. when the societies speaking these languages become more exoteric (*14*), or even disappear, as in Ossetic and Cappadocian Greek, where the loss of gender has also been linked to L2 learning and contact (*15*). Apart from adult L2 speakers failing to faithfully learn a foreign language, the simplification (loss or reduction of phonologically fused marking) can result from L1 speakers consciously or unconsciously accommodating their speech to the needs of the outsiders by reducing grammatical markers that pose acquisition difficulties (*5*).

These links between language and social structure have received a substantial amount of attention, mainly through the lens of focused, small-scale comparisons. A range of qualitative studies have analyzed closely related varieties of Quechua (*8*), English (*11*, *16*), and German (*17*, *18*) as well as Tibeto-Burman (*19*) and Scandinavian languages (*7*). These studies seem to corroborate that, among closely related varieties, the languages that were more exposed to contact with L2 speakers tend to show less irregular, less opaque grammatical markers in the studied domains.

However, the extent to which these findings generalize beyond a handful of cases remains unclear. Each of these studies call on different (and sometimes idiosyncratic) linguistic and sociodemographic variables, which calls into question the homogeneity of the causes and the mechanisms underlying them. This limited comparability was partially addressed by comparative studies that aimed at assessing these hypotheses at a global scale. (*6*) showed that different aspects of morphological complexity are inversely correlated with population size (the number of L1 speakers), geographic spread, and the number of linguistic neighbors, which became known as the Linguistic Niche hypothesis. A follow-up study (*20*) found no correlation between the number of cases in nouns and population size, but showed that this linguistic feature is negatively correlated with the proportion of L2 speakers. Yet other studies (*21*) reveal a negative correlation between verbal synthesis and both demographic variables (the number of L1 speakers and the proportion of L2 speakers) within the same model. Finally, some studies find no relationship between morphological complexity and the presence or absence of a substantial proportion of L2 speakers (*22*). All in all, it remains unclear whether any of these measures of exotericity are meaningfully associated with language structure. Large language samples in these works made it challenging to reliably place the studied communities on an exotericity continuum. Reliable information on all criteria (homogeneity of the population, community size, social network density, relative isolation, etc.) delineating the distinctions between more and less exoteric communities was unavailable, so instead different sociodemographic variables served as proxies for exotericity.

The inconsistent findings of previous studies may have arisen for three reasons. First, the cross-linguistic coverage of these studies varies substantially, and thus raises the question about how representative these samples are of global differences in grammatical complexity. The limited sample size often results from uneven feature coverage in the WALS database (*23*). WALS covers 2,662 languages but only a few hundreds have information available for over 50% of features. This makes studying multiple features associated with complexity impossible without decreasing the sample or having more uncertainty in the data. Second – and perhaps more importantly – the previous studies involve widely different linguistic phenomena that are assumed to be comparable only through the lens of the umbrella term "grammatical complexity". Grammatical complexity has many facets: the number of markers, irregularity, obligatoriness, compositionality, redundancy, and reliance on phonologically fused rather than independent forms (*5*, *6*, *24*, *25*). The multifaceted nature of complexity means that a language is seen as more complex as it increases the number of grammatical cases and determiners, irregular verb forms, non-compositional constructions, agreement patterns, and/or phonologically fused markers expressing different functions. However, different underlying mechanisms can be responsible for the changes of these distinct dimensions of complexity in exoteric societies. For example, when complexity is viewed in terms of compositionality, language structures are claimed to become more compositional (they consist of several interpretable parts rather than of one independently interpretable part) in exoteric societies where high proportions of L2 speakers benefit from additional transparency (*5*). These different dimensions can all change at different rates and under different pressures. Hence, combining several of these dimensions into one metric of grammatical complexity (e.g. (*26*, *22*)) may not shed light on the relationship between grammatical structures and the chosen sociodemographic variables reflecting exotericity.

To test the association between grammatical structures and exotericity of the relevant societies, we introduce metrics that quantify two distinct dimensions of grammatical complexity that have been claimed to be reduced in exoteric societies: (1) the degree of phonologically fused grammatical markers ("fusion"), and (2) the number of obligatory grammatical marking ("informativity"). We obtain grammatical features included in both metrics from the large global dataset Grambank (*27*, *28*) (see Table S1 in Supplementary Materials for the list of Grambank features in both metrics).

The fusion score reflects the degree to which the languages rely on phonologically bound markers (e.g. prefixes and suffixes) as opposed to phonologically independent markers. While phonologically independent markers are independent from other morphemes in their stress and form, phonologically fused markers rely on other morphemes in this respect, which makes the task of acquiring phonologically fused markers by adult L2 learners more difficult. Languages with phonologically fused markers for tense-aspect-mood categories on verbs, case on nouns and pronouns, gender and number agreement, possession, negation, and other features will score higher on this metric. For instance, the highest-scoring languages, Tariana (Arawakan), has overt morphological plural marking on nouns, core and oblique cases on nouns and pronouns, overt morphological marking of mood, aspect distinctions, present, past, and future tenses, morphological passive on verbs, and number and gender agreement on different targets, among others.

The informativity metric offers an insight into the amount of obligatory explicit grammatical distinctions made by languages. The following features, for example, increase informativity: politeness distinction in pronouns, remoteness distinctions in past and future tenses, definite and indefinite articles, and number marking on nouns (singular, dual, plural, trial, paucal; associative plural). Languages will score higher on the informativity metric if their grammars have, for instance, different demonstratives used for visible and non-visible objects, as in Tundra Nenets (Uralic), where *tay°kuy° teda* ‘that reindeer’ would refer to the visible reindeer and *t'exa teda* ‘that reindeer’ would be used when the reindeer is not visible, for example, when it is behind something (*29*). We have not included features in this metric if they concerned distinctions that are generally considered to be universal, such as negation (*30*) and possession (*31*). We are unaware of any language that does not make a distinction between affirmative and negated clauses nor any that lack a productive pattern at all for marking possession. Including such features would not tell us anything meaningful about variation in the expression of grammatical meaning in the world's languages.

Another feature that contributes to informativity is the presence of distinctions between inclusive and exclusive constructions in pronominal systems or verbal indexing. For instance, Māori (Austronesian) disambiguates two potential meanings of ‘we’ left unspecified in English: To produce a sentence *We will go on a walk*, the speakers of Māori obligatorily choose between *tātou* if the interlocutor is joining and *mātou* if the interlocutor is not included in ‘we’ and the speaker goes on a walk with other people (*32*). Apart from excluding the features from the metric that are marked in all or almost all languages, since Grambank was designed to capture features found commonly in the world’s languages, our metric does not contain any extremely rare features marked by only a handful of languages. The informativity metric is not sensitive to whether the information is marked by a fused marker or not.

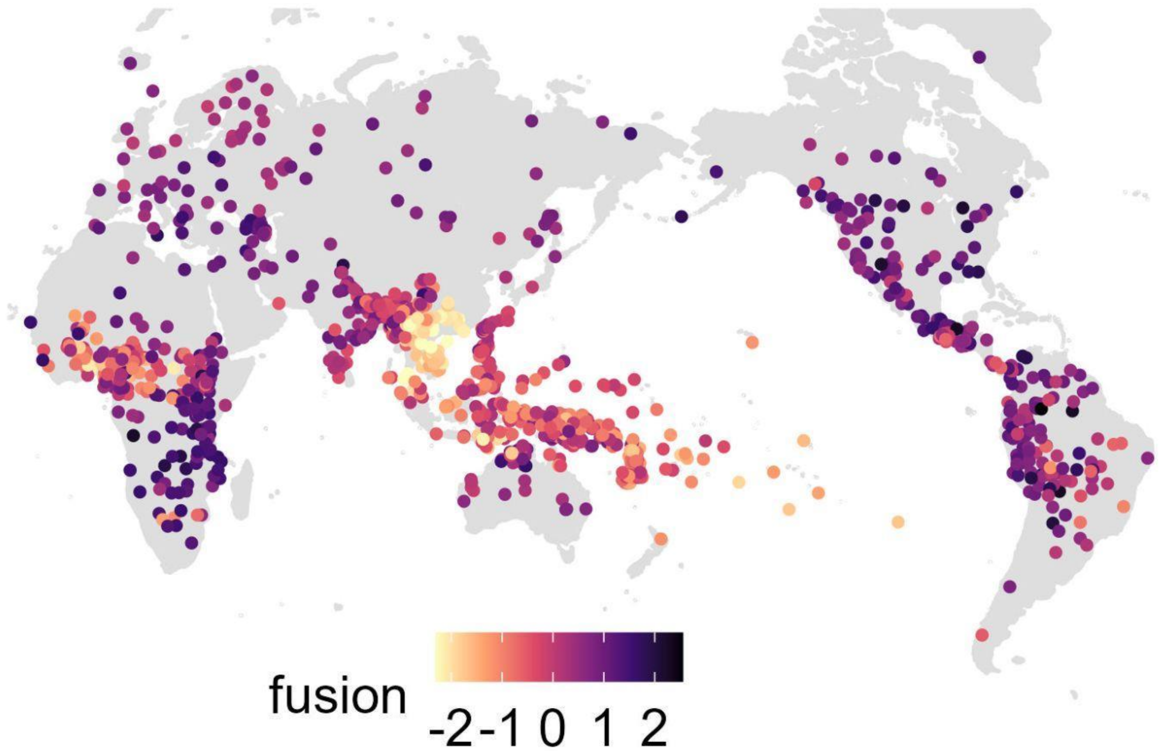
The third potential reason for inconsistent findings in the previous studies is that they do not fully control for phylogenetic and spatial non-independence (but see (*33*)). Specifically, the previous studies control for these confounds in a way that implies assembling languages into large groups based on their ancestry and locations, which oversimplifies the relationships between them. With the exception of (*33*) and (*34*), previous studies tend to treat membership in the same language family as a random effect. However, this approach ignores the relationships between the languages within the same families. Alternatively, the previous studies sample languages from different families and locations. However, this sampling does not always ensure independence of data points (*35*, *36*) and invariably leads to more constrained samples with a subsequent loss in statistical power (*37*). Similarly, random effects of geographical areas such as Glottolog’s macroareas or the 24 AUTOTYP areas (*38*) are used to control for spatial non-independence. With large macroareas, all languages spoken in different continents are grouped together and the differing effect of distance between two neighboring languages and two languages on different sides of the continent are neglected. Random effects of detailed areas have the same problem that individual distances between languages within the same area do not inform the analysis. Additionally, this does not capture the contact between neighboring languages if they belong to two different areas. For example, even though Ukrainian and Polish are closely related geographical neighbors, they are modeled as belonging to distinct areas (Inner Asia and Europe) when AUTOTYP areas are modeled as random effects.

Testing the hypothesis

Here we test the hypothesis that languages in highly exoteric societies have (1) less phonologically fused grammatical markers (fusion) and (2) overall fewer obligatory explicit markers (informativity) compared to languages in low-exotericity societies. We aim to overcome past limitations by (1) using a comprehensive and diverse sample of the world's languages exceeding those in previous studies, (2) motivating the variables involved in the relation between exotericity and language structure, and (3) crafting a state-of-the-art statistical model accounting for the complex historical dependencies between languages.

1. Languages and societies sample

Our sample consists of 1,314 languages (see Figure 1 and Figure 2). The majority of these languages belong to the following language families: Austronesian (296), Sino-Tibetan (144), Atlantic-Congo (140), Afro-Asiatic (58), Austroasiatic (53), and Indo-European (44). Most languages in our sample are located in Oceania (214), Southeast Asia (157), African Savannah (137), and the Indian subcontinent (106). Importantly, Indo-European languages are not overrepresented in our large-scale sample, which is a persistent problem in many cross-linguistic studies.


fusion
-2 -1 0 1 2

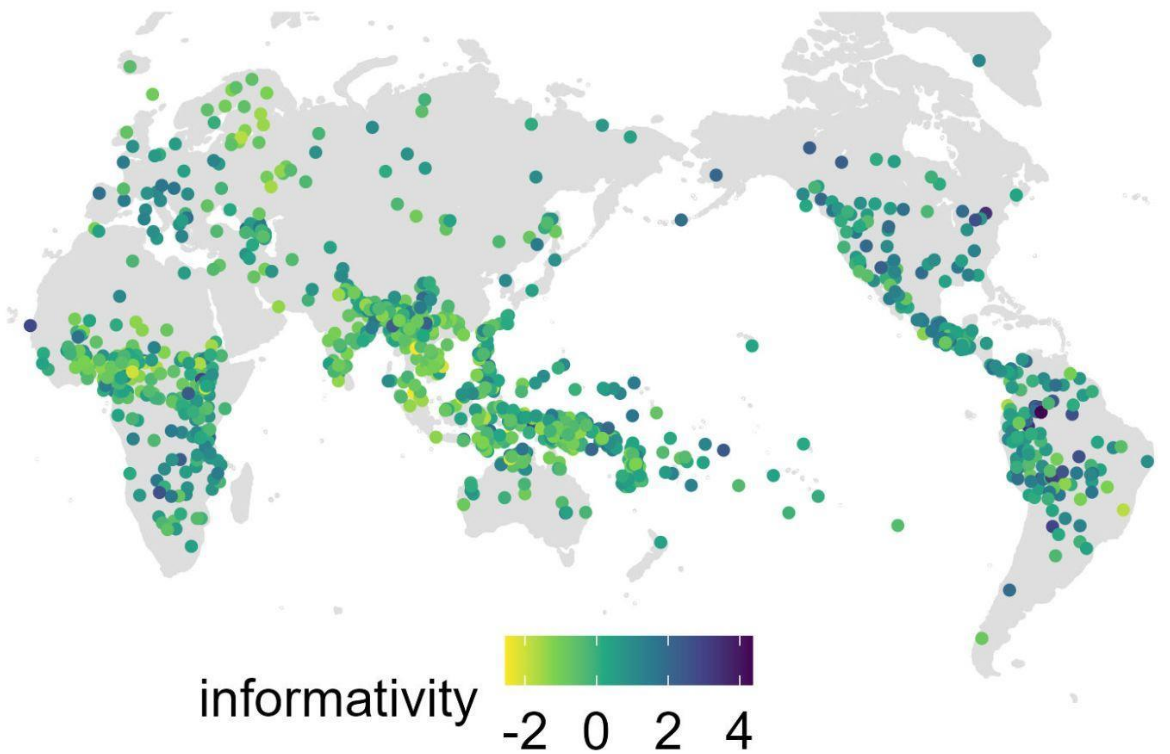

informativity
-2 0 2 4

Figure 1. The global distribution of fusion and informativity scores. The scores with the minimum of 0 (absence of all metric features) and maximum of 1 (presence of all metric features) have been standardized to the mean of 0 and the variance of 1. The hotspots of low fusion are located in West Africa and Southeast Asia. Many Austronesian languages also rank low on fusion. The geographic patterns of informativity scores are less clear compared to fusion. Among lower-scoring languages are those spoken in West Africa, Southeast Asia, many Uralic languages, and languages spoken in India (Indo-Aryan and Dravidian).

2. Sociodemographic variables of exotericity

We examine whether the variation in the scores of fusion and informativity is explained by the sociodemographic factors associated with more exotericity of the societies. This dimension is complex and so far it has escaped simple quantification; however, there are a number of globally available social and demographic variables that correspond to different degrees of exotericity. In general, the society is considered to be more exoteric, if the following metrics are larger or if the binary variables are present:

1) Number of native (L1) speakers (cf. (*6*, *21*));
2) Proportion of non-native (L2) speakers (cf. (*20*, *21*));
3) Number of linguistic neighbors (cf. (*6*));
4) Official status;
5) Usage in education.

The last two variables concerning the status of the language (official/not official; used or not used in education) have not been previously used in predicting grammatical structures. We include these because they help to capture another side of exotericity: exoteric communities are more likely to use languages that are recognized as official languages and languages of education. We predict that these two variables, official status and usage in education, play a crucial role in the Linguistic Niche Hypothesis. They either enable the written form to become more elaborate while the spoken form simplifies (*39*) or they should mitigate the hypothesized negative effect of the number of L1 speakers on grammatical complexity. Both of these factors should act to canonicalise the dominant language, thus preventing the loss of complex features and increasing their transmission fidelity. Additionally, recent studies (*40*) have indicated that language of education in particular is a major cause of minority language loss which will strengthen the selective pressure to learn the dominant language.

All variables are modeled in isolation from each other due to high probability of multicollinearity: a language with many L1 speakers is more likely to have more linguistic neighbors and act as an official language and a language of education.

Even though it has been suggested that larger populations are more likely to have higher proportions of L2 speakers (*6*), it is possible that societies with similar population size might still have different proportions of L2 speakers, which will have different implications for the evolution of these languages if the link between exotericity and grammatical structures holds true. For instance, a society with a large L1 speaker population and few L2 speakers will be lower on the exotericity scale than a language with a similar population size but an extreme proportion of L2 speakers. Due to this and the importance of accounting for multiple linguistic and social factors (*21*, *41*), we additionally fit two models that use the number of L1 speakers and proportion of L2 speakers as 1) two separate fixed effects and 2) as an interaction term between them.

3. Spatiophylogenetic modeling

Using a Bayesian phylogenetic framework, we map fusion and informativity scores obtained from Grambank with available information about the locations from Glottolog 4.4 (*42*) and the sociodemographic variables to the EDGE tree (*43*). As a result, we have a global sample of 1,314 languages available on the phylogeny for which both metric scores were calculated and sociodemographic data was present.

We adopt spatiophylogenetic modeling (*44*) that allows us to study the relation between sociodemographic and linguistic factors while taking into account the complex spatial and genealogical relations between languages and societies. Both spatial and genealogical relations are represented as random effects built on the basis of covariance matrices that stand for the relevant historical processes.

We first fit different combinations of random effects to determine whether the distribution of fusion and informativity scores depends on phylogenetic and geographical dependencies of the languages. In this step, we build seven models containing the intercept and random effects as predictors:

1) Phylogenetic effects
2) Spatial effects: "local" diffusion of scores across several hundreds of kilometers is possible
3) Spatial effects: "regional" diffusion of scores across distances up to 1,000 kilometers (see Methods and Materials section and "Spatial effects" subsection and Figure S1 in Supplementary Materials for more details)
4) Spatial effects: 24 language areas from AUTOTYP
5) Phylogenetic effects + Spatial effects: local
6) Phylogenetic effects + Spatial effects: regional
7) Spatial effects: 24 language areas from AUTOTYP

Secondly, we choose the strongest model to test whether adding sociodemographic variables to it will improve its fit. This implies fitting seven models with different sociodemographic variables or their combinations treated as fixed effects:

1) Number of L1 speakers
2) Proportion of L2 speakers
3) Number of L1 speakers and proportion of L2 speakers (combined model) (cf. (*21*))
4) The interaction term between these two variables (number of L1 speakers * proportion of L2 speakers)
5) Number of linguistic neighbors
6) Official status (binary)
7) Language of education (binary)

We fit these seven models without any random effects to compare to which extent controlling for non-independence influences the results. Then, we compare the models of fusion and informativity to determine the influential predictors of the metric scores. We compare all competing models in our analyses based on obtained Widely Applicable Information Criterion values (WAIC, (*45*, *46*)).

Results

Out of the set of random effects models, both fusion and informativity scores are best predicted by the combination of phylogenetic and spatial effects (Tables 1 & 2). The spatiophylogenetic models incorporating both random effects substantially outperform other models, in particular, the runner-up phylogenetic-only models. This indicates that both effects explain variation in the scores better than phylogenetic effects in isolation. The differences between WAIC values (*45*) between the strongest spatiophylogenetic models and other models are larger than 45 for both fusion and informativity. The preference for the “local” as opposed to the “regional” version of the spatial random effect suggests the likely diffusion of the scores across short distances of several hundreds of kilometers. Whereby the random effects assuming a much wider possible diffusion, i.e. “regional” spatial effects (>1,000 kilometers) or the random effects of 24 language areas, performed worse.

This best-fitting model incorporating phylogenetic and spatial effects is then used to test whether adding any of the sociodemographic predictors (or their combinations) contributed to understanding the distribution of fusion and informativity. We find that the effects of these predictors range from negligible to low. The strongest models predicting fusion and informativity are those incorporating these random effects and the number of L1 speakers. The second-best models additionally include the proportion of L2 speakers. The models including other social variables did not outperform the spatiophylogenetic models.

However, the linear regression coefficients of most fixed effects, including the proportion of L2 speakers, are negligible: they either overlap with or are not appreciably different from zero (see Figure 2). Only one of the tested sociodemographic variables is a robust, yet weak predictor of metrics scores. For both fusion and informativity, we find a weak positive correlation with the number of L1 speakers variable in two models where it is 1) the only sociodemographic variable and 2) in combination with proportion of L2 speakers (the 95% Credible Intervals for fusion and informativity: 0.04-0.14 and 0.05-0.2 respectively). Importantly, none of these relationships are negative as predicted by prior studies. This contradicts the argument advanced by the Linguistic Niche Hypothesis that there should be an inverse relationship between grammatical complexity and the sociodemographic factors associated with exotericity.

All models with fixed effects that exclude random effects of phylogenetic and spatial relationships rank lower than the same models that additionally implement random effects (see Table S2 in Supplementary Materials). This means that predictive performance of models without random effects is inferior compared to the models that incorporate both fixed and random effects.

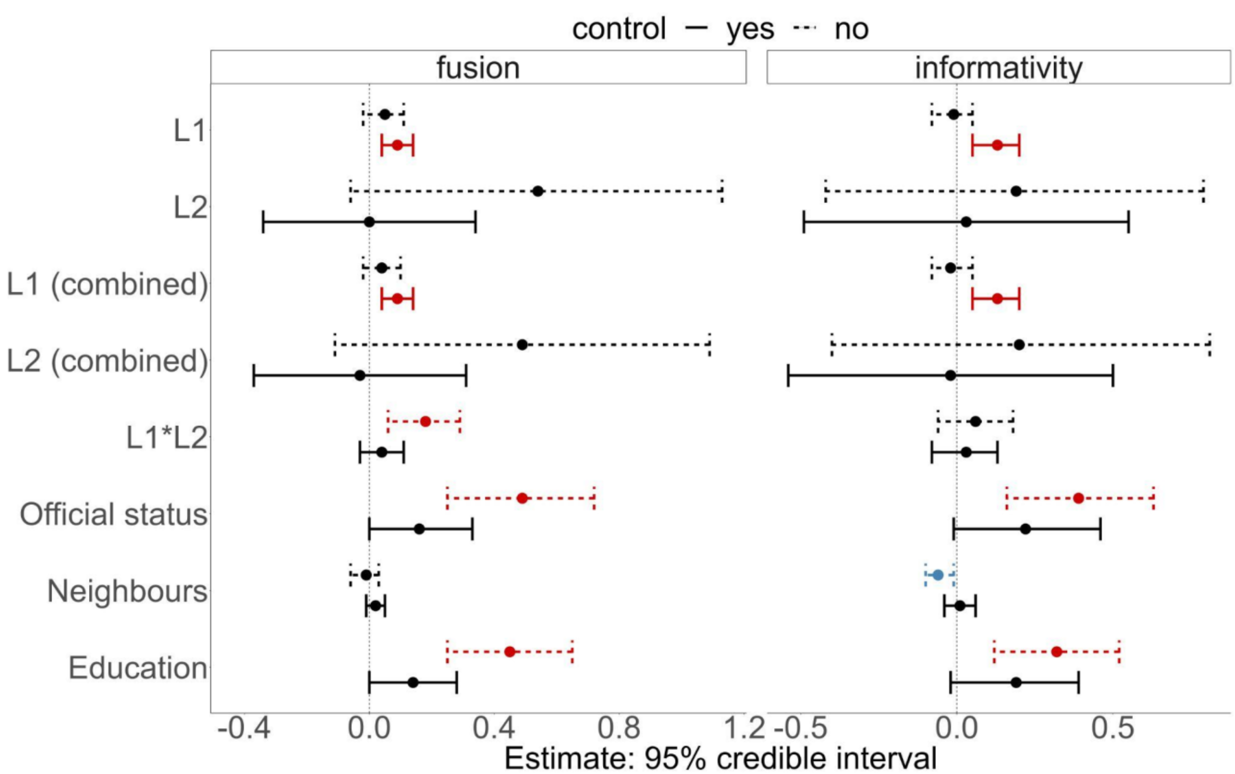


Figure 2. The coefficients and 95% credible intervals for fixed effects in 6 bivariate regression models and one multivariable model (L1 combined & L2 combined) for fusion and informativity, without and with spatiophylogenetic random effects (dashed and solid lines respectively). The linear regression coefficients of fixed effects representing exotericity in the spatiophylogentic models are depicted with the error bars. Error bars in black cross zero, whereas the bars colored red and blue indicate robust positive and negative relationships. All effects that appear influential (colored in red or blue) in models that do not control for random effects of genealogy and geography (dashed error bars) disappear after we control for these sources of non-independence (solid error bars). The exceptions are the weak positive effects of L1 speakers on fusion and

informativity that are revealed in the full model with fixed and random effects but remain hidden in the model where the number of L1 speakers is the only predictor of metric scores. This shows that controlling for non-independence of languages is indispensable for unraveling the proposed relationships of dependence between grammatical structures and sociodemographic factors. The short names of the effects of the number of L1 speakers and the proportion of L2 speakers differ based on the model they are incorporated in: the models where these variables are modeled in isolation (L1 (log-transformed and standardized number of L1 speakers) and L2 effects), the combined model with both of these effects (L1 (combined) (log-transformed and standardized number of L1 speakers) and L2 (combined)), and the model with an interaction term (L1*L2), where the log-transformed number of L1 speakers was used).

Instead, the distribution in fusion and informativity scores is largely explained by phylogenetic random effects – 92% and 70% of the total variance of fusion and informativity – and spatial random effects that account for 4% and 8% of the total variance of the scores correspondingly (see Table S3 in Supplementary Materials). We measure the phylogenetic signal of both complexity dimensions on the global tree by estimating Pagel's lambda (λ) (*47*). The values of this metric range from 0, which indicates no phylogenetic signal (random distribution of scores with respect to the phylogeny), to 1, which indicates strong phylogenetic signal (closely related languages share similar scores), or greater than 1. Both fusion and informativity show strong phylogenetic signal: phylogenetic signal of fusion (λ=0.97, <0.001) is stronger than that of informativity (λ=0.85, <0.001). In other words, fusion and informativity scores are explained by the inheritance from a common ancestor and spatial diffusion amongst close neighbors. Nevertheless, the role of spatial relationships might have been downplayed as the information about locations of languages has informed the structure of the global EDGE tree, in that weak geographic priors were imposed on likely language relationships within the phylogeographic model. This means that the relative contribution of the spatial predictor might be larger.

One alternative explanation for why we find no substantial effect of exotericity on fusion and informativity is that these relationships are nonlinear. It might be the case, for instance, if the effects of the number of L1 speakers or the proportion of L2 speakers on grammatical structures are only substantial for extremely large or extremely small communities. To address this possibility, apart from fitting linear regressions with the number of L1 speakers and proportion of L2 speakers, we operationalize them in the form of nonlinear effects within the random walk models of order 2 (RW2). We find that these models rank similarly to their counterparts with corresponding linear effects. Since the standard deviation of these random nonlinear effects is small (<0.08 for fusion and informativity), we report only the results of linear regression models in the main text and provide Table S2 summarizing the WAIC values and effects of all fitted models in Supplementary Materials.

Discussion

The specific claim of the "Linguistic Niche hypothesis" that grammatical complexity should reduce with increased number of non-native speakers is not supported by our results. Contrary to the expected inverse correlation between complexity scores and sociodemographic variables reflecting exotericity, the only robust relationships we find are weak positive effects of L1 speakers on fusion and informativity.

Instead, we found that the two dimensions of complexity we model – fusion and informativity – are better predicted by genealogy and geographic diffusion than by exotericity measures. Measuring the phylogenetic signal of these two features also showed that the distribution of their scores was largely influenced by the shared evolutionary histories between languages on the global tree. Both of these dimensions of grammatical complexity appear to be highly stable phylogenetically (see Figure 3), which suggests that fusion and informativity are phylogenetically constrained.

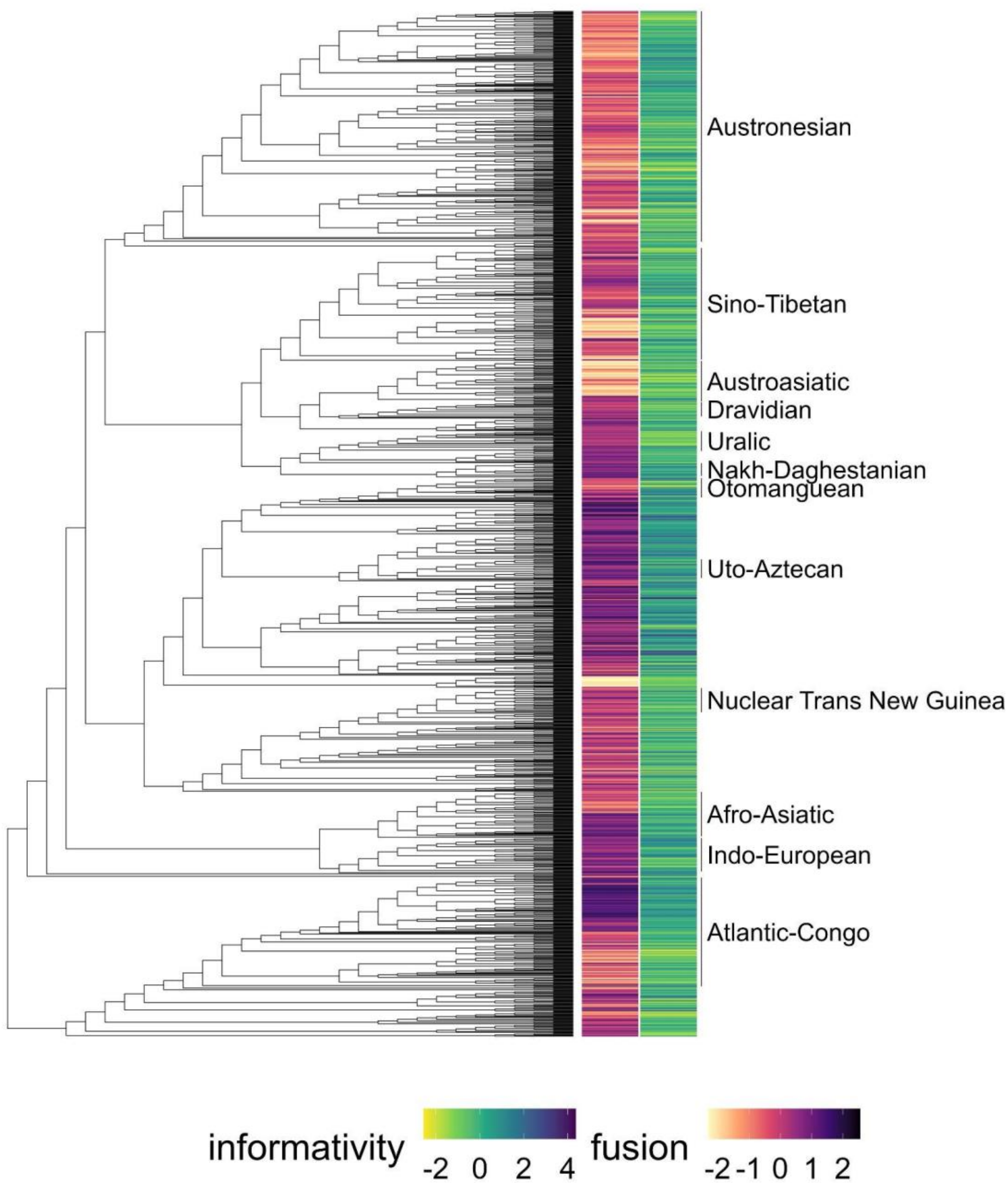


Figure 3. The scores of fusion and informativity on the global tree. We detect many patterns of closely-related languages scoring similarly, which might indicate the faithful transmission of grammatical complexity from ancestor languages to their descendants rather than large-scale adaptations of grammatical complexity to changes in sociodemographic factors. Similar to geographic distribution, we see that fusion scores follow a more defined pattern of phylogenetic clustering compared to informativity scores.

Taking a closer look at the languages spoken in Southern Africa shows why phylogenetic distances explain the variation in fusion scores better than demographic properties of the societies (see Figure 4.A). Most languages in this area have high fusion scores, including the high-scoring language, Southern Sotho (1.29): a Southern Bantu language used in a highly exoteric society by > 5.5 million L1 speakers and > 7 million L2 speakers. By contrast, Tsonga, another Southern Bantu language, scores notably low: -0.71. It is also spoken in an exoteric niche, but its degree of exotericity is lower since it has less L2 speakers (> 3 million) than Southern Sotho, even though it has approximately 1 million more L1 speakers. Given their sociolinguistic environments, we could expect more resemblance in the scores between these two languages and more fusion in Tsonga. However, we observe a less pronounced difference between Tsonga's score and that of its neighboring sister language Tswa with an above-average (0.58) score, spoken in a more esoteric niche by > 6 million L1 speakers and no L2 speakers. Some other languages with notably low fusion scores in Southern Africa are languages from other language families: East Taa (Tuu) (-1.35) with 2,500 L1 speakers and severely endangered Amkoe (Kxa) (-1.1) with several dozen L1 speakers. Despite being spoken in highly esoteric societies, these languages rank extremely low on fusion. The fact that out of the examined languages in Southern Africa the lower scoring ones (or moderately scoring ones as is the case for Tswa) are either two closely related Bantu languages or languages from language families other than Bantu indicates that genealogical relationships and to some extent potential contact between neighboring languages provide a better basis for understanding the variation in fusion scores.

The Uralic family exemplifies a case where phylogenetic effects might be challenging to disentangle from spatial ones because the distribution of informativity scores can be explained by both phylogenetic relationships and geographic distances/contact phenomena. Most Uralic languages score low on the informativity metric (see Figure 4.B1-2). The exceptions with higher scores are languages belonging to branches such as Samoyedic (Nganasan, Selkup, Tundra Nenets, and Forest Enets), Ugric (Hungarian), and Permian (Komi-Permyak and Komi-Zyrian), which diverged earlier from the rest of the languages subsumed under the Mari, Mordvin, Saami, and Finnic branches (*46*). The contrasts in informativity scores between these sister languages do not seem to match with the hypothesis that lower informativity should exist in exoteric societies. For instance, Votic with 25 speakers scores lower in informativity (-1.79) than its closest relative Estonian (-0.8) with >1 million speakers. Even though phylogenetic distances serve as an explanation for the distribution of the scores in the family, geographic proximity/contact may have been another influential factor. For instance, Hungarian matches the scores of the surrounding higher-scoring Indo-European languages. Similarly, Tundra Nenets and Komi-Permyak, might score high in informativity due to the bilingualism of their speakers in Russian. Unlike higher proportions of adult L2 speakers, child bilingualism has been associated with an increase of grammatical complexity rather than its loss (*11*).

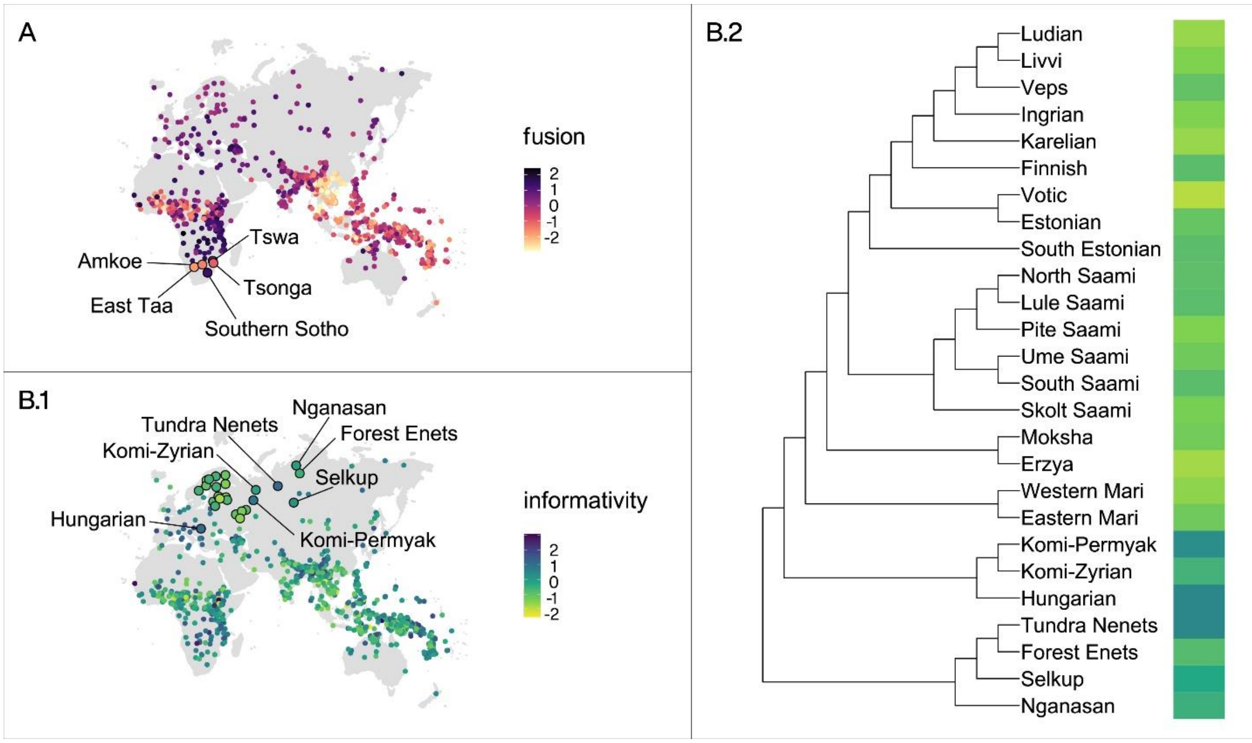


Figure 4. The distribution of fusion scores in Southern Africa (**A**) and informativity scores in Eurasia (**B.1**) with the focus on Uralic languages and the phylogeny of Uralic languages included in the global tree (**B.2**). The difference in fusion scores between two Bantu languages – Tsonga (low) and Southern Sotho (high) – can be explained from the perspective of phylogenetic relatedness, with Tsonga resembling the scores of other low-scoring outliers in Southern Africa from other language families than Bantu and to lesser extent its neighboring sister language Tswa (**A**). Higher informativity scores in Uralic languages show a clear pattern of phylogenetic clustering (**B.2**) and are found in Samoyedic (Nganasan, Selkup, Tundra Nenets, and Forest Enets), Ugric (Hungarian), and Permian (Komi-Permyak and Komi-Zyrian), which diverged earlier from the rest of the languages. Higher informativity scores in Uralic languages can also be ascribed to contact with Indo-European languages (**B.1**): Hungarian is surrounded by higher-scoring languages spoken in Europe and the speakers of Tundra Nenets and Komi-Permyak are typically bilingual in Russian.

One of the key differences between the analysis presented here and previous studies is that we use spatiophylogenetic methods to explicitly model the effects of genealogical and geographic non-independence. This allows us not only to address two sources of non-independence between languages in our cross-linguistic samples (*40*, *48*) but also to quantify and compare the relative importance of the phylogenetic and spatial effects. The majority of previous studies ignore the dependence between languages belonging to the same family or located in the same area by treating families and areas as random effects, or sampling languages from distinct families and locations in the attempt to exclude languages that are non-independent (*35*, *36*). Our results clearly show that this methodological difference really matters (Figure 2). Only the number of L1 speakers emerges as a weak effect after controlling for genealogy and geography, whereas most other effects appear influential only in models without these random effects and disappear when the non-independence of languages is controlled for. In some cases, variables like fusion/informativity scores and the proportion of L2 speakers are not associated in both models that do and do not account for genealogy and geography. This lack of the association can instead be attributed to our large sample and how we quantified two complexity dimensions.

Another advantage of our approach is the use of two complexity metrics on a comprehensive global sample. Previous empirical studies either focused on one grammatical domain, such as the number of cases (*20*) or verbal synthesis (*21*), or they included a wide variety of features corresponding to different coding procedures and interpretations of complexity (see studies based on WALS (*23*), such as (*6*, *22*, *26*)). Moreover, we followed a systematic focused approach toward delineating fusion and informativity, which allows us to make principled decisions about the choice of the Grambank features for each metric and avoid including features that did not align with either of the described interpretations. Further studies could explore whether other grammatical complexity dimensions are sensitive to the influence of sociodemographic factors. One promising avenue for future research would be measuring the degree to which languages deviate from the principle "one-meaning–one-form" (*25*) on a cross-linguistic scale.

All in all, the previous positive findings might be artifacts of small non-representative samples, metrics subsuming grammatical features falling under distinct complexity dimensions or methods not sufficiently controlling for genealogy and geography. Having overcome these past limitations and used a large sample offering greater statistical power, we find no evidence that societies of strangers speak grammatically simpler languages and that sociodemographic factors used in this study are strong drivers of fusion and informativity. Future studies could build on the methodological approach adopted in this paper and examine the impact of a more nuanced set of sociohistorical variables. The global Grambank dataset could be used to study other ambitious questions about linguistic diversity and language evolution.

While other language structures (e.g. lexicon (*5*, *49*, *50*) or individual grammatical features e.g. case marking (*20*, *41*) might adapt to changing sociodemographic factors, we find no evidence that the two dimensions of grammatical complexity measured here respond to sociolinguistic pressures in the hypothesised manner. We find that phylogenetic inheritance and borrowing between near neighbors explain most of the distribution of grammatical complexity among the world's languages. This finding suggests that, even if speaker population size does play a role in driving down language complexity, the strength of selection is weak. Population size can change rapidly and unpredictably due to external events (wars, diseases, and migrations), let alone natural population growth (*51*). These changes might not leave noticeable traces on grammatical complexity for two reasons: Either the phylogenetically stable nature of these complexity variables constrains the rate at which these traits can adapt to their sociolinguistic environment, or the lability of population size means adaptation lags behind the selection pressures from the new selective regime.

Future work should explore the potential effects of other sociodemographic factors that are more fine-grained than the currently available demographic variables. Attention should be devoted to those factors that change at a relatively slow pace and hence provide relatively enduring selection pressures. Future studies should carefully consider the interplay between genealogy and geography when modelling the adaptation of languages to their environments and explain why some features might be more sensitive than others to sociolinguistic pressures.

Materials and Methods

Datasets

Metrics. In this study, we want to focus on two delineated interpretations of grammatical complexity to better understand the forces shaping them:
1) the degree of fusion (the extent to which the language relies on phonologically fused markers (*52*))
2) the number of explicit and obligatory grammatical distinctions not routinely marked in all languages.

These linguistic phenomena are estimated by two metrics which are based on features available in Grambank 1.0. The first metric measures fusion. It accounts for phonologically fused grammatical markers, with scores increasing for having more phonologically fused markers. For each Grambank feature that concerns fused marking, a language can get 1 or 0.5 “fusion scores” if it is coded as “present” in the database. If for a given feature, the language does not use the fused marker (is coded as “absent”) it receives 0 fusion scores. We then take the mean of all these fusion scores per language to construct the score.

For instance, for plural marking on nouns, a language like English that forms plural forms with phonologically fused markers (-*s*) gets 1 fusion score for this feature. By contrast, languages like Rapa Nui or Māori that lack phonologically fused markers for the purpose of nominal plural marking, or Vietnamese with no plural marking on nouns, are assigned 0 for this fusion feature. There are 6 features that cover multiple different kinds of morphological marking, including additive morphology such as affixes but also other morphological processes like suppletion and tone which are not typically included as indications of fusion in studies of grammatical complexity. These features receive a fusion score of 0.5, since they do not meet the necessary condition for being classified as fused markers in all languages where they are present but could represent fused marking in some of these languages (unlike for example word-order features which cannot indicate fused marking at all). These include productive patterns of derivation (e.g. a productive morphological pattern for deriving an action/state noun from a verb), morphological cases on pronouns, and a morpho-syntactic distinction between predicates expressing controlled versus uncontrolled events or states. These features are assigned a fusion score of 0.5 to avoid treating them as features exclusively referring to phonologically fused markers while at the same time not treating them as being entirely unrelated to fused marking. Overall, this metric systematically targets phonologically fused grammatical markers and can be compared to other metrics that capture the degree of morphological complexity (*6*), inventory complexity (*53*), and syntheticity (*16*).

The fusion metric is built on Grambank data, which in turn relies on reading grammatical descriptions and communication with experts. We acknowledge that different authors may have different approaches to what they define as "fused" and "phonologically independent". However, despite these potential differences, the resulting fusion scores are in agreement with how previous works rank the world's languages in their morphological complexity (see Materials and Methods section for the comparison of fusion scores and the scores and judgements of morphological complexity in previous literature).

The second metric consists of features that contribute to explicit obligatory marking of grammatical and semantic distinctions that are not routinely marked in all languages. To ensure this, we exclude the typically overt grammatical domains related to negation, possession, comparative constructions, polar interrogation, and marking of A and P arguments (with the help of word order, case on nominal words, or indexes on verbs). Languages get scores for the presence of informativity features. Phonologically fused and independent markers contribute equally to the final informativity score when they address the same grammatical function. This means that if a language has an obligatory marking of plurality with phonologically fused and/or phonologically independent markers, it will get assigned 1 informativity score. Using the previous example, English, Māori, and Rapa Nui would all be assigned 1, while Vietnamese would have 0. This metric reveals the degree of grammatical marking in languages and is similar in spirit to measuring grammaticity (*16*)) with a caveat that we exclude certain domains that are usually explicitly and obligatorily marked in most languages (negation, polar interrogation etc.).

Next, the scores for each grammatical function in both metrics are summed, and the mean value is obtained for each language, so that the possible minimum for a language that has no fusion-related or no informative features is 0, and the possible maximum value is 1, if a language has all features that count as fusion-related or informative. In reality, we find no languages that reach the maximum scores on either of the metrics, but various languages (almost) approach the minimum of 0. For instance, one of the lowest-scoring (0) languages on fusion is Hu (Austroasiatic) lacking all fusion features, while the lowest informativity score 0.1 is obtained by Jukun-Takum (Atlantic-Congo). Tariana (Arawakan) reaches the highest scores in both metrics: 0.65 on fusion and 0.66 on informativity. Next, the scores ranging between 0 and 1 are standardized to the mean of 0 and the variance of 1.

We measure fusion and informativity only for those languages that are well-described in Grambank (*27*) and remove languages with over 25% of missing values across all Grambank features. Out of 2,430 languages in the dataset, we compute the fusion and informativity scores for 1,362 languages. This way the resulting scores are robust representations of the targeted complexity dimensions.

Importantly, despite our focus on phonologically fused markers rather than morphological inflections, our metric of fusion is still comparable with what other studies measured as morphological complexity. For instance, according to the metric based on a variety of WALS features, Turkish ranked extremely high (0.775) and Vietnamese (Austroasiatic) was the lowest-scoring language (0.141) (*26*). In our data, we do not quantify fusion scores for Vietnamese, but observe a divide between Turkish and some Austroasiatic languages lacking all fusion features: Turkish is assigned 0.43 with the maximally bound language Tariana (Arawakan) reaching as high as 0.65, while the fusion score of Thavung (same Vietic branch as Vietnamese), Hu, Prai, and Rumai Palaung is equal 0. Similarly, our metric captures the contrasts between languages from Kiranti and Kuki-Chin branches as suggested in (*19*). Camling (Sino-Tibetan) is claimed to show extreme morphological complexity and scores 0.42 in our metric, while Mara Chin (Sino-Tibetan) is said to have lost complexity, which is reflected in its lower score: 0.14 (*19*). Furthermore, the geographical patterns in the distribution of fusion is in line with typological literature and complexity studies. One prominent hotspot of low fusion is located in Mainland Southeast Asia, which is expected based on the typological profiles of languages in this area (*54*), and another one is in West Africa, which is in line with the proposal of the Low-Complexity-Belt (*55*).

Sociodemographic variables. Fusion and informativity are predicted based on the following demographic and social variables: number of L1 speakers and proportion of L2 speakers (*56*); number of linguistic neighbors, the status of the language (official/not official), and usage in education (language of education/not language of education), obtained from supplementary information available in (*40*), with data gathered by the authors or originally retrieved from World Language Mapping System v17 and v16 (WLMS, http://worldgeodatasets.com) and (*57*).

We log-transform the raw numbers of L1 speakers with a base of 10 and then standardize this variable along with the number of linguistic neighbors to have a mean of 0 and variance of 1. Standardization is done to eliminate the potential effects of extreme outliers on the results of the spatiophylogenetic modeling. The log-transformed (but not standardized) number of L1 speakers is used for implementing the interaction term between this variable and L2 proportion.

The number of linguistic neighbors was calculated as the number of intersections between a 10,000 km^2 circle of a given language and the polygons of other languages (*40*). The polygons and single-point locations were obtained from the World Language Mapping System (WLMS e16 and e17, http://worldgeodatasets.com); single-point locations were used to approximate language areas using Voronoi projections where WLMS provided no polygons (*40*).

The status of the language is a binary variable: the language status is either not official or official at the national level, and the language is either used or not used in education. The languages are coded as official if they are formally recognized as such or if they are treated as the main languages of education, commerce, media, and government in countries that do not formally recognize any language as official, such as Australia (*40*). In such cases, these languages are also coded as languages of education if no other language serves as the language of education in the territory (*40*). All minority languages that were not recognized as official but were used as media of instruction in education according to *L'aménagement linguistique dans le monde* (*57*) were also assigned to be languages of education (see (*40*) and accompanying materials for more details on coding of the used sociodemographic variables).

We acknowledge that the modeling of grammatical features with fluctuating sociodemographic predictors poses challenges. We note that 1) the data from Grambank and the sources of sociodemographic variables are typical of contemporary populations and 2) the sociodemographic variables are particularly prone to change more rapidly (*33*, *58*). Previous studies interpret the correlation between synchronic grammatical and sociodemographic structures diachronically. However, it is not clear how much time is necessary for the changes in the sociolinguistic environment to set in motion the changes to grammatical structures. Moreover, the use of the common demographic variables, such as population size, proportion of L2 speakers, and the number of linguistic neighbors, as proxies for exotericity has been questioned in a number of studies (*21*, *34*, *59*, *60*). While these limitations are inherent to all cross-linguistic studies on the link between grammatical structures and exotericity, our study overcomes past limitations concerning the sample size, control for sources of non-independence, and the choice of linguistic features.

Phylogenetic and geographic information. We use Glottolog 4.4 (*42*) for information on the locations and language family of the languages in our sample. The AUTOTYP database (*38*) provided information on the distribution of languages across 24 language areas.
To model the evolutionary changes in these features over time and control for shared ancestry, we map grammatical complexity scores and values of sociodemographic variables onto the global EDGE (evolutionarily distinct, globally endangered) phylogeny (*43*). This global supertree integrates language classification information from Glottolog and published phylogenies, as well as their locations.

Spatiophylogenetic modelling

We adopt a spatiophylogenetic modeling technique pioneered by (*44*). This Bayesian approach uses an Integrated Nested Laplace Approximation (INLA) (*61*, *62*) to estimate the joint posterior distribution of model parameters and is implemented in *R* (*63*) package *INLA* (*61*). On top of evaluating fixed effects coefficients, spatiophylogenetic modeling allows us to calculate the relative influence of random effects – here, spatial (geographic distances) and phylogenetic relationships – on the response variable.

Building random effects. To incorporate phylogenetic and spatial relationships as random effects in the models, we undertake several steps to represent these relationships in the form of precision matrices as required within INLA. We make these matrices comparable by standardizing them to have a variance of 1. This is done in the following way.

We build a phylogenetic variance-covariance matrix which quantifies the shared branch lengths between languages on the global tree with the assumption of a Brownian motion model of evolution using *vcv.phylo* function in *ape* (*64*). The tree serves to build a standardized phylogenetic precision matrix by applying *inverseA* function in *MCMCglmm* package (*65*).

We follow a similar set of steps to calculate and standardized spatial matrices. First, we calculate a variance-covariance matrix under the Matérn spatial covariance function implemented in *varcov.spatial* function in *geoR* package (*66*). Second, we standardize the variance-covariance matrix by the variance, and invert it to create a precision matrix. Finally, the standardized variance-covariance matrix is transformed into the precision matrix. We estimate two spatial matrices for two sets of parameters: a) ϕ=1.25 and κ=1 ("local set") and b) ϕ=17 and κ=1 ("regional set") (see (*67*) and (*28*) for other examples of how this type of control for spatial autocorrelation is implemented). These parameters of the spatial covariance matrix were chosen to differentiate between two assumptions: under parameters corresponding to the local set, the diffusion of similar metric scores between languages are not likely across distances over 1,000 km, while with the regional set parameters the diffusion can take place over several thousands of kilometers (see Figure S1 in Supplementary Materials for details). Fitting the models with each of these matrices allows us to compare which assumption about the diffusion of metrics scores corresponds to our data: are languages more likely to have similar scores only locally across hundreds of kilometers or is diffusion likely across larger regions, such as continents or large language areas spanning thousands of kilometers. For comparison, we also introduce the third control for spatial autocorrelation: random effects of 24 areas from the AUTOTYP database (*38*), where each area is treated as independent from each other, whereby we neglect geographic distances between the areas. These three different spatial effects are seen in competition with each other and represent different ways of operationalizing the influence of contact between neighboring languages. The models incorporating phylogenetic effects and “local” versions of spatial effects predict fusion and informativity best (see Table 1 and Table 2).

Sensitivity testing. We use Penalized Complexity (PC) priors (*68*) on the precision of the likelihood and phylogenetic and spatial effects. Precision matrices are standardized to have a variance of 1. In the main text results, PC priors are set so that 10% of the prior probability density of the standard deviation of the likelihood or random effects fall above 1. As a sensitivity test, we vary the probability density at 1%, 10%, 50%, and 99%, but this does not affect our conclusions (see Table S4 in Supplementary Materials).

Measuring phylogenetic signal. We estimate Pagel’s lambda (λ) (*47*) to measure phylogenetic signal of fusion and informativity on the global EDGE tree using *R* package *phytools* (*69*). The values of λ typically range from 0 to 1. λ=1 implies a high phylogenetic signal, which means that the scores evolve in a manner expected under a Brownian motion model. Conversely, λ=0 suggests no phylogenetic signal and indicates that the distribution of scores evolved independently from the phylogenetic relationships between the languages in the phylogeny.

Table 1. WAIC values and quantiles (0.025, 0.5, and 0.975) of estimates of models fitting only random effects and intercept in models predicting fusion.

model	effect	2.5%	50%	97.5%	WAIC
Phylogenetic+Spatial: local+L1 speakers	**phylogenetic SD**	1.48	1.71	1.99	1,822.66
	spatial SD	0.29	0.35	0.42	
	Intercept	-0.03	-0.01	0.02	
	L1	0.04	0.09	0.14	
Phylogenetic+Spatial: local+L1 speakers+L2 proportion	**phylogenetic SD**	1.48	1.71	1.99	1,824.81
	spatial SD	0.29	0.35	0.42	
	Intercept	-0.03	-0.01	0.02	
	L1	0.04	0.09	0.14	
	L2 proportion	-0.37	-0.03	0.31	

Phylogenetic+Spatial: local+Neighbours	**phylogenetic SD**	1.47	1.70	1.97	1,849.87
	spatial SD	0.29	0.35	0.42	
	Intercept	-0.02	0.00	0.02	
	Neighbours	-0.01	0.02	0.05	
Phylogenetic+Spatial: local+L1_log10:L2 proportion	**phylogenetic SD**	1.46	1.69	1.96	1,853.72
	spatial SD	0.29	0.35	0.42	
	Intercept	-0.02	0.00	0.02	
	L1*L2 proportion	-0.03	0.04	0.11	
Phylogenetic+Spatial: local+L2 proportion	**phylogenetic SD**	1.47	1.70	1.97	1,854.33
	spatial SD	0.29	0.35	0.42	
	Intercept	-0.02	0.00	0.03	
	L2 proportion	-0.34	0.00	0.34	
Phylogenetic+Spatial: local+Official	**phylogenetic SD**	1.45	1.67	1.95	1,854.91
	spatial SD	0.29	0.36	0.42	
	Intercept	-0.03	-0.01	0.02	
	Official status	0.00	0.16	0.33	
Phylogenetic+Spatial: local+Education	**phylogenetic SD**	1.45	1.68	1.95	1,855.62
	spatial SD	0.29	0.35	0.42	
	Intercept	-0.03	-0.01	0.02	
	Education	0.00	0.14	0.28	

Table 2. WAIC values and quantiles (0.025, 0.5, and 0.975) of estimates of models fitting only random effects and intercept in models predicting informativity.

model	effect	2.5%	50%	97.5%	WAIC
Phylogenetic+Spatial: local+L1 speakers	**phylogenetic SD**	0.97	1.28	1.65	3,207.14

	spatial SD	0.35	0.43	0.53	
	Intercept	-0.03	0.01	0.05	
	L1	0.05	0.13	0.20	
Phylogenetic+Spatial: local+L1 speakers+L2 proportion	**phylogenetic SD**	0.97	1.28	1.65	3,209.11
	spatial SD	0.35	0.43	0.53	
	Intercept	-0.03	0.01	0.05	
	L1	0.05	0.13	0.20	
	L2 proportion	-0.54	-0.02	0.50	
Phylogenetic+Spatial: local+Neighbours	**phylogenetic SD**	0.88	1.20	1.57	3,229.76
	spatial SD	0.36	0.44	0.54	
	Intercept	-0.01	0.03	0.07	
	Neighbours	-0.04	0.01	0.06	
Phylogenetic+Spatial: local+L2 proportion	**phylogenetic SD**	0.88	1.20	1.56	3,229.95
	spatial SD	0.36	0.44	0.54	
	Intercept	-0.01	0.03	0.07	
	L2 proportion	-0.49	0.03	0.55	
Phylogenetic+Spatial: local+L1_log10:L2 proportion	**phylogenetic SD**	0.88	1.19	1.56	3,230.07
	spatial SD	0.36	0.45	0.54	
	Intercept	-0.01	0.03	0.07	
	L1*L2 proportion	-0.08	0.03	0.13	
Phylogenetic+Spatial: local+Education	**phylogenetic SD**	0.85	1.16	1.53	3,233.24
	spatial SD	0.36	0.44	0.54	
	Intercept	-0.03	0.01	0.06	
	Education	-0.02	0.19	0.39	

Phylogenetic+Spatial: local+Official	**phylogenetic SD**	0.83	1.14	1.52	3,233.87
	spatial SD	0.37	0.45	0.54	
	Intercept	-0.03	0.02	0.06	
	Official status	-0.01	0.22	0.46	

References

1. W. R. Thurston, *Processes of change in the languages of North-Western New Britain* (Pacific Linguistics, Canberra, 1987).

2. W. R. Thurston, "How exoteric languages build a lexicon: esoterogeny in West New Britain" in *VICAL 1: Oceanic Languages. Papers from the Fifth International Conference on Austronesian Linguistics, Auckland, New Zealand, January 1988*, R. Harlow, R. Hooper, Eds. (Auckland, New Zealand, 1989), vol. 1, pp. 555–579.

3. T. Givón, P. Young, Cooperation and interpersonal manipulation in the society of intimates. *Typological Studies in Language*. **48**, 23–56 (2002).

4. T. Givón, *Context as other minds: The pragmatics of sociality, cognition, and communication* (John Benjamins Publishing, 2005), vol. 130.

5. A. Wray, G. W. Grace, The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*. **117**, 543–578 (2007).

6. G. Lupyan, R. Dale, Language Structure Is Partly Determined by Social Structure. *PLOS ONE*. **5**, e8559 (2010).

7. W. Kusters, *Linguistic complexity* (Netherlands Graduate School of Linguistics, 2003).

8. W. Kusters, Complexity in linguistic theory, language learning and language change. *Language complexity: Typology, contact, change*. **94**, 3–22 (2008).

9. W. R. Thurston, "Sociolinguistic typology and other factors effecting change in northwestern New Britain, Papua New Guinea" in *Culture change, language change: Case studies from Melanesia*, T. E. Dutton, Ed. (Pacific Linguistics. (Pacific Linguistics C-120), Canberra, 1992), pp. 123–139.

10. A. Aikhenvald, *A grammar of Tariana, from northwest Amazonia* (Cambridge University Press, Cambridge, 2003).

11. P. Trudgill, *Sociolinguistic typology: Social determinants of linguistic complexity* (Oxford University Press, Oxford, 2011).

12. H. Clahsen, C. Felser, K. Neubauer, M. Sato, R. Silva, Morphological structure in native and nonnative language processing. *Language learning*. **60**, 21–43 (2010).

13. R. Dale, G. Lupyan, Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in complex systems*. **15**, 1150017 (2012).

14. Ö. Dahl, *The growth and maintenance of linguistic complexity* (John Benjamins Publishing, Amsterdam, 2004).

15. I. Igartua, Loss of grammatical gender and language contact. *Diachronica*. **36**, 181–221 (2019).

16. B. Szmrecsanyi, B. Kortmann, "Between simplification and complexification: Non-standard varieties of English around the world" in *Language complexity as an evolving variable*, G. Sampson, D. Gil, P. Trudgill, Eds. (Oxford University Press, Oxford, 2009), pp. 64–79.

17. P. Maitz, A. Németh, Language contact and morphosyntactic complexity: Evidence from German. *Journal of Germanic Linguistics*. **26**, 1–29 (2014).

18. R. Baechler, "Diachronic complexification and isolation" in *Yearbook of the Poznan Linguistic Meeting* (2014), vol. 1, pp. 1–28.

19. S. DeLancey, The Historical Dynamics of Morphological Complexity in Trans-Himalayan. *Linguistic Discovery*. **13** (2015).

20. C. Bentz, B. Winter, Languages with more second language learners tend to lose nominal case. *Quantifying language dynamics*, 96–124 (2013).

21. K. Sinnemäki, F. Di Garbo, Language structures may adapt to the sociolinguistic environment, but it matters what and how you count: A typological study of verbal and nominal complexity. *Frontiers in psychology*. **9**, 1141 (2018).

22. A. Koplenig, Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *Royal Society open science*. **6**, 181274 (2019).

23. M. Dryer, M. Haspelmath, Eds., The World Atlas of Language Structures Online (2013), (available at http://wals.info).

24. J. McWhorter, *Language interrupted: Signs of non-native acquisition in standard language grammars* (Oxford University Press, Oxford, 2007).

25. M. Miestamo, "Grammatical complexity in a cross-linguistic perspective" in *Language complexity: Typology, contact, change*, M. Miestamo, K. Sinnemäki, F. Karlsson, Eds. (John Benjamins, Amsterdam/Philadelphia, 2008), pp. 23–41.

26. C. Bentz, T. Ruzsics, A. Koplenig, T. Samardzic, "A comparison between morphological complexity measures: typological data vs. language corpora" in *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity, Osaka, Japan, December 11-17 2016* (2016), pp. 142–153.

27. The Grambank Consortium (eds.), Grambank 1.0. (2022), (available at http://grambank.clld.org).

28. H. Skirgård, H. J. Haynie, D. E. Blasi, H. Hammarström, J. Collins, J. Latarche, J. Lesage, T. Weber, A. Witzlack-Makarevich, S. Passmore, A. Chira, L. Maurits, R. Dinnage, M. Dunn, G. Reesink, R. Singer, C. Bowern, P. Epps, J. Hill, O. Vesakoski, M. Robbeets, K. Abbas, D. Auer, N. Bakker, G. Barbos, R. Borges, S. Danielsen, L. Dorenbusch, E. Dorn, J. Elliott, G. Falcone, J. Fischer, Y. Ghanggo Ate, H. Gibson, H. Göbel, J. Goodall, V. Gruner, A. Harvey, R. Hayes, L. Heer, R. Herrera Miranda, N. Hübler, B. Huntington-Rainey, J. Ivani, M. Johns, E. Just, E. Kashima, C. Kipf, J. Klingenberg, N. König, K. Koti, R. Kowalik, O. Krasnoukhova, N. Lindvall, M. Lorenzen, H. Lutzenberger, T. Martins, C. Mata German, S. van der Meer, J. Montoya Samamé, M. Müller, S. Muradoglu, K. Neely, J. Nickel, M. Norvik, C. A. Oluoch, J. Peacock, I. Pearey, N. Peck, S. Petit, S. Pieper, M. Poblete, D. Prestipino, L. Raabe, A. Raja, J. Reimringer, S. Rey, J. Rizaew, E. Ruppert, K. Salmon, J. Sammet, R. Schembri, L. Schlabbach, F. Schmidt, A. Skilton, W. D. Smith, H. de Sousa, K. Sverredal, D. Valle, J. Vera, J. Voß, T. Witte, H. Wu, S. Yam, J. Ye 葉婧婷, M. Yong, T. Yuditha, R. Zariquiey, R. Forkel, N. Evans, S. C. Levinson, M. Haspelmath, S. J. Greenhill, Q. D. Atkinson, R. D. Gray, Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss (2022), , doi:10.31235/osf.io/grh45.

29. I. Nikolaeva, *A Grammar of Tundra Nenets* (Berlin: Mouton de Gruyter, 2014), vol. 65 of *Mouton Grammar Library*.

30. H. Zeijlstra, Negation in natural language: On the form and meaning of negative elements. *Language and Linguistics Compass*. **1**, 498–518 (2007).

31. J. Nichols, B. Bickel, M. S. Dryer, M. Haspelmath, Eds., Locus of Marking in Possessive Noun Phrases (v2020.3). *The World Atlas of Language Structures Online* (2013), , doi:10.5281/zenodo.7385533.

32. R. Harlow, *Māori* (Lincom Europa, 1996), vol. 20.

33. C. Bentz, A. Verkerk, D. Kiela, F. Hill, P. Buttery, Adaptive Communication: Languages with More Non-Native Speakers Tend to Have Fewer Word Forms. *PLOS ONE*. **10**, e0128254 (2015).

34. A. Verkerk, F. Di Garbo, Sociogeographic correlates of typological variation in northwestern Bantu gender systems. *Language Dynamics and Change*. **12**, 155–223 (2022).

35. E. A. Eff, Does Mr. Galton still have a problem? Autocorrelation in the standard cross-cultural sample. *World Cultures*. **15**, 153–170 (2004).

36. L. Bromham, Curiously the same: swapping tools between linguistics and evolutionary biology. *Biology & Philosophy*. **32**, 855–886 (2017).

37. S. C. Levinson, R. D. Gray, Tools from evolutionary biology shed new light on the diversification of languages. *Trends in cognitive sciences*. **16**, 167–173 (2012).

38. B. Bickel, J. Nichols, T. Zakharko, A. Witzlack-Makarevich, K. Hildebrandt, M. Rießler, L. Bierkandt, F. Zúñiga, J. B. Lowe, The AUTOTYP database (v1.1.0) (2022), , doi:10.5281/zenodo.6793367.

39. G. Lupyan, R. Dale, Why are there different languages? The role of adaptation in linguistic diversity. *Trends in cognitive sciences*. **20**, 649–660 (2016).

40. L. Bromham, R. Dinnage, H. Skirgård, A. Ritchie, M. Cardillo, F. Meakins, S. Greenhill, X. Hua, Global predictors of language endangerment and the future of linguistic diversity. *Nature ecology & evolution*. **6**, 163–173 (2022).

41. K. Sinnemäki, Linguistic system and sociolinguistic environment as competing factors in linguistic variation: A typological approach. *Journal of Historical Sociolinguistics*. **6**, 20191010 (2020).

42. H. Hammarström, R. Forkel, M. Haspelmath, S. Bank, Glottolog 4.4 (2021), (available at https://doi.org/10.5281/zenodo.4761960).

43. R. Bouckaert, D. Redding, O. Sheehan, T. Kyritsis, R. Gray, K. E. Jones, Q. Atkinson, Global language diversification is linked to socio-ecology and threat status (2022) (available at https://osf.io/preprints/socarxiv/f8tr6/).

44. R. Dinnage, A. Skeels, M. Cardillo, Spatiophylogenetic modelling of extinction risk reveals evolutionary distinctiveness and brief flowering period as threats in a hotspot plant genus. *Proceedings of the Royal Society B*. **287**, 20192817 (2020).

45. S. Watanabe, Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of machine learning research*. **11**, 3571–3594 (2010).

46. A. Gelman, J. Hwang, A. Vehtari, Understanding predictive information criteria for Bayesian models. *Statistics and computing*. **24**, 997–1016 (2014).

47. M. Pagel, Inferring the historical patterns of biological evolution. *Nature*. **401**, 877–884 (1999).

48. R. Mace, M. Pagel, The comparative method in anthropology. *Current Anthropology*. **35**, 549–564 (1994).

49. L. Raviv, A. Meyer, S. Lev-Ari, Compositional structure can emerge without generational transmission. *Cognition*. **182**, 151–164 (2019).

50. L. Raviv, A. Meyer, S. Lev-Ari, Larger communities create more systematic languages. *Proceedings of the Royal Society B*. **286**, 20191262 (2019).

51. S. J. Greenhill, X. Hua, C. F. Welsh, H. Schneemann, L. Bromham, Population size and the rate of language evolution: a test across Indo-European, Austronesian, and Bantu languages. *Frontiers in Psychology*. **9**, 576 (2018).

52. B. Bickel, J. Nichols, Inflectional morphology. *Language typology and syntactic description*. **3**, 169–240 (2007).

53. J. Nichols, C. Bentz, "Morphological complexity of languages reflects the settlement history of the Americas" in *New Perspectives on the Peopling of the Americas*, K. Harvati, G. Jäger, H. Reyes-Centeno, Eds. (Kerns Verlag, Tübingen, 2018), pp. 13–26.

54. N. J. Enfield, Areal linguistics and mainland Southeast Asia. *Annu. Rev. Anthropol.* **34**, 181–206 (2005).

55. C. Bentz, "The Low-Complexity-Belt: Evidence for large-scale language contact in human prehistory" in *Proceedings of the 11th International Conference (EVOLANG11) on The Evolution of Language, New Orleans, LA, USA, 21 March 2016* (2016).

56. D. M. Eberhard, G. F. Simons, C. D. Fennig, Eds., Ethnologue: Languages of the World (2020), (available at www.ethnologue.com).

57. J. Leclerc, L'aménagement linguistique dans le monde (2019), (available at https://www.axl.cefan.ulaval.ca/monde/index_alphabetique.htm).

58. K. Sinnemäki, "Complexity in core argument marking and population size" in *Language complexity as an evolving variable*, G. Sampson, D. Gil, P. Trudgill, Eds. (Oxford University Press, Oxford, 2009), pp. 126–140.

59. F. Di Garbo, E. Kashima, R. Napoleão de Souza, K. Sinnemäki, Concepts and methods for integrating language typology and sociolinguistics. *Tipologia e Sociolinguistica: Verso un approccio integrato allo studio della variazione: Atti del Workshop della Società Linguistica Italiana 20 settembre 2020*. **5**, 143–176 (2021).

60. F. Di Garbo, A. Verkerk, A typology of northwestern Bantu gender systems. *Linguistics*. **60**, 1169–1239 (2022).

61. H. Rue, S. Martino, N. Chopin, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*. **71**, 319–392 (2009).

62. T. G. Martins, D. Simpson, F. Lindgren, H. Rue, Bayesian computing with INLA: new features. *Computational Statistics & Data Analysis*. **67**, 68–83 (2013).

63. R. C. Team, others, R: A language and environment for statistical computing (2022).

64. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. **35**, 526–528 (2019).

65. J. D. Hadfield, MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*. **33**, 1–22 (2010).

66. P. J. Ribeiro, P. J. Diggle, M. Schlather, R. Bivand, B. Ripley, *geoR: Analysis of Geostatistical Data* (2020; https://CRAN.R-project.org/package=geoR).

67. S. Claessens, T. Kyritsis, Q. D. Atkinson, Revised analysis shows relational mobility predicts sacrificial behavior in Footbridge but not Switch or Loop trolley problems. *Proceedings of the National Academy of Sciences*. **117**, 13203–13204 (2020).

68. D. Simpson, H. Rue, A. Riebler, T. G. Martins, S. H. Sørbye, Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical science*. **32**, 1–28 (2017).

69. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other things). *Methods in ecology and evolution*. **3**, 217–223 (2012).

Acknowledgments

We thank the reviewers and editors for their comments. We are grateful for the aid in the technical application of INLA by Russell Dinnage. We appreciate the help in graphic design by Hans-Georg Sell, in particular, with Figure 4. D.E.B. was supported by a

Branco Weiss Fellowship and a Harvard Data Science Fellowship. We thank the audiences of *55th Annual Meeting of the Societas Linguistica Europaea* and *Joint Conference on Language Evolution* for the feedback.

Funding:
Department of Linguistic and Cultural Evolution (OS, RDG, SJG, HS)
Branco Weiss Fellowship (DEB)
Harvard Data Science Fellowship (DEB)

Author contributions:

Conceptualization: OS, HS, SMM, HJH
Metrics design: HS, HJH
Metrics revision: OS, SMM
Methodology: OS, HS, SP, HJH, SJG, DEB
Software: OS, SP, HS, SJG
Formal analysis and investigation: OS
Visualization: OS, HS, SP, SJG
Supervision: HS, SMM, DEB, SJG, RDG, VG
Writing—original draft: OS, SMM, DEB, HS, RDG
Writing—review & editing: all authors

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All data, code, and materials are available in the main text, the Supplementary Materials, or a public repository and the scientific archive Zenodo, with the exception of:

- the entire Grambank dataset: as this database is not currently publicly available, the output file with scores of two grammatical complexity metrics rather than the entire dataset with all features will be provided on the public repository.
- the entire Ethnologue dataset in its original form: the variables from Ethnologue (standardized number of L1 speakers, standardized and log-transformed number of L1 speakers, standardized number of all language users, standardized and log-transformed number of all language users, and proportion of L2 speakers) are made publicly available under a completed material transfer agreement. The access to these variables in the public repository allows fitting all models in the analyses except for the model including the interaction between log-transformed number of L1 speakers and the proportion of L2 speakers. The agreement allows to make only those variables publicly available that cannot be transformed back to their original form. Since log-transformed number of L1 speakers could be transformed back, a special set of scripts within the repository has been created for the users without the access to Ethnologue in order to run all models except for the mentioned one.

Supplementary Materials

Supplementary materials include Tables S1 to S4, Supplementary Text (Section S1 Spatial effects), and Figure S1.

Supplementary Materials for

Societies of strangers do not speak grammatically simpler languages

Olena Shcherbakova[*] *et al.*

*Corresponding author. Email: shcherbakova@shh.mpg.de

Table S1. The overview of Grambank features in metrics of fusion and informativity.

Feature_ID	Feature	fusion	informativity
GB074	Are there prepositions?	0	
GB075	Are there postpositions?	0	
GB303	Is there a phonologically free antipassive marker ("particle" or "auxiliary")?	0	antipassive
GB520	Can aspect be marked by a non-inflecting word ("auxiliary particle")?	0	aspect
GB519	Can mood be marked by a non-inflecting word ("auxiliary particle")?	0	mood
GB299	Can standard negation be marked by a non-inflecting word ("auxiliary particle")?	0	
GB317	Is dual number regularly marked in the noun phrase by a phonologically free element?	0	dual
GB320	Is paucal number regularly marked in the noun phrase by a phonologically free element?	0	paucal
GB318	Is plural number regularly marked in the noun phrase by a phonologically free element?	0	plural
GB316	Is singular number regularly marked in the noun phrase by a phonologically free element?	0	singular
GB319	Is trial number regularly marked in the noun phrase by a phonologically free element?	0	trial
GB302	Is there a phonologically free passive marker ("particle" or "auxiliary")?	0	passive

GB262	Is there a clause-initial polar interrogative particle?	0	
GB263	Is there a clause-final polar interrogative particle?	0	
GB264	Is there a polar interrogative particle that most commonly occurs neither clause-initially nor clause-finally?	0	
GB521	Can tense be marked by a non-inflecting word ("auxiliary particle")?	0	tense
GB071	Are there morphological cases for independent personal pronominal core arguments (i.e. S/A/P)?	0.5	
GB146	Is there a morpho-syntactic distinction between predicates expressing controlled versus uncontrolled events or states?	0.5	control
GB047	Is there a productive morphological pattern for deriving an action/state noun from a verb?	0.5	
GB048	Is there a productive morphological pattern for deriving an agent noun from a verb?	0.5	
GB049	Is there a productive morphological pattern for deriving an object noun from a verb?	0.5	
GB073	Are there morphological cases for independent oblique personal pronominal arguments (i.e. not S/A/P)?	0.5	
GB148	Is there a morphological antipassive marked on the lexical verb?	1	antipassive
GB070	Are there morphological cases for non-pronominal core arguments (i.e. S/A/P)?	1	

GB091	Can the A argument be indexed by a suffix/enclitic on the verb in the simple independent clause?	1	
GB092	Can the A argument be indexed by a prefix/proclitic on the verb in the simple independent clause?	1	
GB093	Can the P argument be indexed by a suffix/enclitic on the verb in the simple independent clause?	1	
GB094	Can the P argument be indexed by a prefix/proclitic on the verb in the simple independent clause?	1	
GB086	Is a morphological distinction between perfective and imperfective aspect available on verbs?	1	aspect
GB120	Can aspect be marked by an inflecting word ("auxiliary verb")?	1	aspect
GB188	Is there any productive augmentative marking on the noun (exclude marking by system of nominal classification only)?	1	augmentative
GB275	Is there a bound comparative degree marker on the property word in a comparative construction?	1	
GB187	Is there any productive diminutive marking on the noun (exclude marking by system of nominal classification only)?	1	diminutive
GB170	Can an adnominal property word agree with the noun in gender/noun class?	1	
GB172	Can an article agree with the noun in gender/noun class?	1	

GB198	Can an adnominal numeral agree with the noun in gender/noun class?	1	
GB119	Can mood be marked by an inflecting word ("auxiliary verb")?	1	mood
GB312	Is there overt morphological marking on the verb dedicated to mood?	1	mood
GB107	Can standard negation be marked by an affix, clitic or modification of the verb?	1	
GB298	Can standard negation be marked by an inflecting word ("auxiliary verb")?	1	
GB171	Can an adnominal demonstrative agree with the noun in gender/noun class?	1	
GB184	Can an adnominal property word agree with the noun in number?	1	
GB185	Can an adnominal demonstrative agree with the noun in number?	1	
GB186	Can an article agree with the noun in number?	1	
GB043	Is there productive morphological dual marking on nouns?	1	dual
GB166	Is there productive morphological paucal marking on nouns?	1	paucal
GB044	Is there productive morphological plural marking on nouns?	1	plural
GB042	Is there productive overt morphological singular marking on nouns?	1	singular
GB165	Is there productive morphological trial marking on nouns?	1	trial

GB147	Is there a morphological passive marked on the lexical verb?	1	passive
GB285	Can polar interrogation be marked by a question particle and verbal morphology?	1	
GB286	Can polar interrogation be indicated by overt verbal morphology only?	1	
GB430	Can adnominal possession be marked by a prefix on the possessor?	1	
GB431	Can adnominal possession be marked by a prefix on the possessed noun?	1	
GB432	Can adnominal possession be marked by a suffix on the possessor?	1	
GB433	Can adnominal possession be marked by a suffix on the possessed noun?	1	
GB115	Is there a phonologically bound reciprocal marker on the verb?	1	reciprocity
GB114	Is there a phonologically bound reflexive marker on the verb?	1	reflexivity
GB121	Can tense be marked by an inflecting word ("auxiliary verb")?	1	tense
GB082	Is there overt morphological marking of present tense on verbs?	1	tense
GB083	Is there overt morphological marking on the verb dedicated to past tense?	1	tense
GB084	Is there overt morphological marking on the verb dedicated to future tense?	1	tense
GB103	Is there a benefactive applicative marker on the verb (including indexing)?	1	benefactive

GB108	Is there directional or locative morphological marking on verbs?	1	directional
GB104	Is there an instrumental applicative marker on the verb (including indexing)?	1	instrumental
GB149	Is there a morphologically marked inverse on verbs?	1	inverse
GB152	Is there a morphologically marked distinction between simultaneous and sequential clauses?	1	simultanseq
GB151	Is there an overt verb marker dedicated to signalling coreference or noncoreference between the subject of one clause and an argument of an adjacent clause ("switch reference")?	1	switch reference
GB072	Are there morphological cases for oblique non-pronominal NPs (i.e. not S/A/P)?	1	
GB079	Do verbs have prefixes/proclitics, other than those that only mark A, S or P (do include portmanteau: A & S + TAM)?	1	
GB080	Do verbs have suffixes/enclitics, other than those that only mark A, S or P (do include portmanteau: A & S + TAM)?	1	
GB081	Is there productive infixation in verbs?	1	
GB089	Can the S argument be indexed by a suffix/enclitic on the verb in the simple independent clause?	1	
GB090	Can the S argument be indexed by a prefix/proclitic on the verb in the simple independent clause?	1	
GB113	Are there verbal affixes or clitics that turn intransitive verbs into transitive ones?	1	

GB155	Are causatives formed by affixes or clitics on verbs?	1	
GB177	Can the verb carry a marker of animacy of argument, unrelated to any gender/noun class of the argument visible in the NP domain?	1	argumentanimacy
GB059	Is the adnominal possessive construction different for alienable and inalienable nouns?	not included	alienability
GB057	Are there numeral classifiers?	not included	numera classifers
GB058	Are there possessive classifiers?	not included	possessive classifiers
GB333	Is there a decimal numeral system?	not included	
GB334	Is there synchronic evidence for any element of a quinary numeral system?	not included	
GB335	Is there synchronic evidence for any element of a vigesimal numeral system?	not included	
GB336	Is there a body-part tallying system?	not included	
GB022	Are there prenominal articles?	not included	
GB023	Are there postnominal articles?	not included	
GB314	Can augmentative meaning be expressed productively by a shift of gender/noun class?	not included	augmentative
GB276	Is there a non-bound comparative degree marker modifying the property word in a comparative construction?	not included	

GB265	Is there a comparative construction that includes a form that elsewhere means 'surpass, exceed'?	not included	
GB266	Is there a comparative construction that employs a marker of the standard which elsewhere has a locational meaning?	not included	
GB270	Can comparatives be expressed using two conjoined clauses?	not included	
GB273	Is there a comparative construction with a standard marker that elsewhere has neither a locational meaning nor a 'surpass/exceed' meaning?	not included	
GB421	Is there a preposed complementizer in complements of verbs of thinking and/or knowing?	not included	
GB422	Is there a postposed complementizer in complements of verbs of thinking and/or knowing?	not included	
GB315	Can diminutive meaning be expressed productively by a shift of gender/noun class?	not included	diminutive
GB031	Is there a dual or unit augmented form (in addition to plural or augmented) for all person categories in the pronoun system?	not included	pronoundualaug
GB136	Is the order of core argument (i.e. S/A/P) constituents fixed?	not included	
GB408	Is there any accusative alignment of flagging?	not included	
GB409	Is there any ergative alignment of flagging?	not included	
GB410	Is there any neutral alignment of flagging?	not included	

GB053	Is there a gender/noun class system where animacy is a factor in class assignment?	not included	genderanimacy
GB054	Is there a gender/noun class system where plant status is a factor in class assignment?	not included	genderplant
GB051	Is there a gender/noun class system where sex is a factor in class assignment?	not included	gendersex
GB052	Is there a gender/noun class system where shape is a factor in class assignment?	not included	gendershape
GB197	Is there a male/female distinction in 1st person independent pronouns?	not included	pronoungender1
GB196	Is there a male/female distinction in 2nd person independent pronouns?	not included	pronoungender2
GB030	Is there a gender distinction in independent 3rd person pronouns?	not included	pronoungender3
GB137	Can standard negation be marked clause-finally?	not included	
GB138	Can standard negation be marked clause-initially?	not included	
GB039	Is there nonphonological allomorphy of noun number markers?	not included	
GB041	Are there several nouns (more than three) which are suppletive for number?	not included	
GB025	What is the order of adnominal demonstrative and noun?	not included	
GB024	What is the order of numeral and noun in the NP?	not included	

GB065	What is the pragmatically unmarked order of adnominal possessor noun and possessed noun?	not included	
GB193	What is the order of adnominal property word and noun?	not included	
GB203	What is the order of the adnominal collective universal quantifier ('all') and the noun?	not included	
GB130	What is the pragmatically unmarked order of S and V in intransitive clauses?	not included	
GB131	Is a pragmatically unmarked constituent order verb-initial for transitive clauses?	not included	
GB132	Is a pragmatically unmarked constituent order verb-medial for transitive clauses?	not included	
GB133	Is a pragmatically unmarked constituent order verb-final for transitive clauses?	not included	
GB257	Can polar interrogation be marked by intonation only?	not included	
GB260	Can polar interrogation be indicated by a special word order?	not included	
GB291	Can polar interrogation be marked by tone?	not included	
GB297	Can polar interrogation be indicated by a V-not-V construction?	not included	
GB250	Can predicative possession be expressed with a transitive 'habeo' verb?	not included	
GB252	Can predicative possession be expressed with an S-like possessum and a locative-coded possessor?	not included	

GB253	Can predicative possession be expressed with an S-like possessum and a dative-coded possessor?	not included	
GB254	Can predicative possession be expressed with an S-like possessum and a possessor that is coded like an adnominal possessor?	not included	
GB256	Can predicative possession be expressed with an S-like possessor and a possessum that is coded like a comitative argument?	not included	
GB026	Can adnominal property words occur discontinuously?	not included	
GB306	Is there a phonologically independent non-bipartite reciprocal pronoun?	not included	reciprocity
GB305	Is there a phonologically independent reflexive pronoun?	not included	reflexivity
GB327	Can the relative clause follow the noun?	not included	
GB328	Can the relative clause precede the noun?	not included	
GB046	Is there an associative plural marker for nouns?	not included	assocplural
GB028	Is there a distinction between inclusive and exclusive?	not included	clusivity
GB027	Are nominal conjunction and comitative expressed by different elements?	not included	comitative
GB117	Is there a copula for predicate nominals?	not included	copulaprednom
GB325	Is there a count/mass distinction in interrogative quantifiers?	not included	count_mass
GB020	Are there definite or specific articles?	not included	definitearticles

GB035	Are there three or more distance contrasts in demonstratives?	not included	demonstartivedistance
GB038	Are there demonstrative classifiers?	not included	demonstrative classifiers
GB036	Do demonstratives show an elevation distinction?	not included	demonstrativeelevation
GB037	Do demonstratives show a visible-nonvisible distinction?	not included	demonstrativevisibility
GB140	Is verbal predication marked by the same negator as all of the following types of predication: locational, existential and nominal?	not included	differentneg
GB322	Is there grammatical marking of direct evidence (perceived with the senses)?	not included	evidentiality_direct
GB323	Is there grammatical marking of indirect evidence (hearsay, inference, etc.)?	not included	evidentiality_indirect
GB126	Is there an existential verb?	not included	existentialverb
GB021	Do indefinite nominals commonly have indefinite articles?	not included	indef
GB309	Are there multiple past or multiple future tenses, distinguishing distance from Time of Reference?	not included	multipletense
GB415	Is there a politeness distinction in 2nd person forms?	not included	politeness
GB127	Are different posture verbs used obligatorily depending on an inanimate locatum's shape or position (e.g. 'to lie' vs. 'to stand')?	not included	postureverbs

GB139	Is there a difference between imperative (prohibitive) and declarative negation constructions?	not included	prohibitive
GB167	Is there a logophoric pronoun?	not included	pronounlog
GB116	Do verbs classify the shape, size or consistency of absolutive arguments by means of incorporated nouns, verbal affixes or suppletive verb stems?	not included	verbclassify
GB068	Do core adjectives (defined semantically as property concepts such as value, shape, age, dimension) act like verbs in predicative position?	not included	
GB069	Do core adjectives (defined semantically as property concepts; value, shape, age, dimension) used attributively require the same morphological treatment as verbs?	not included	
GB095	Are variations in marking strategies of core participants based on TAM distinctions?	not included	
GB096	Are variations in marking strategies of core participants based on verb classes?	not included	
GB098	Are variations in marking strategies of core participants based on person distinctions?	not included	
GB099	Can verb stems alter according to the person of a core participant?	not included	
GB105	Can the recipient in a ditransitive construction be marked like the monotransitive patient?	not included	
GB109	Is there verb suppletion for participant number?	not included	

GB110	Is there verb suppletion for tense or aspect?	not included	
GB111	Are there conjugation classes?	not included	
GB118	Are there serial verb constructions?	not included	
GB122	Is verb compounding a regular process?	not included	
GB123	Are there verb-adjunct (aka light-verb) constructions?	not included	
GB124	Is incorporation of nouns into verbs a productive intransitivizing process?	not included	
GB129	Is there a notably small number, i.e. about 100 or less, of verb roots in the language?	not included	
GB134	Is the order of constituents the same in main and subordinate clauses?	not included	
GB135	Do clausal objects usually occur in the same position as nominal objects?	not included	
GB150	Is there clause chaining?	not included	
GB156	Is there a causative construction involving an element that is unmistakably grammaticalized from a verb for 'to say'?	not included	
GB158	Are verbs reduplicated?	not included	
GB159	Are nouns reduplicated?	not included	
GB160	Are elements apart from verbs or nouns reduplicated?	not included	

GB192	Is there a gender system where a noun's phonological properties are a factor in class assignment?	not included	
GB204	Do collective ('all') and distributive ('every') universal quantifiers differ in their forms or their syntactic positions?	not included	
GB296	Is there a phonologically or morphosyntactically definable class of ideophones that includes ideophones depicting imagery beyond sound?	not included	
GB300	Does the verb for 'give' have suppletive verb forms?	not included	
GB301	Is there an inclusory construction?	not included	
GB304	Can the agent be expressed overtly in a passive clause?	not included	
GB313	Are there special adnominal possessive pronouns that are not formed by an otherwise regular process?	not included	
GB321	Is there a large class of nouns whose gender/noun class is not phonologically or semantically predictable?	not included	
GB324	Is there an interrogative verb for content interrogatives (who?, what?, etc.)?	not included	
GB326	Do (nominal) content interrogatives normally or frequently occur in situ?	not included	
GB329	Are there internally-headed relative clauses?	not included	
GB330	Are there correlative relative clauses?	not included	
GB331	Are there adjoined relative clauses?	not included	

GB400	Are all person categories neutralized in some voice, tense, aspect, mood and/or negation?	not included	
GB401	Is there a class of patient-labile verbs?	not included	
GB402	Does the verb for 'see' have suppletive verb forms?	not included	
GB403	Does the verb for 'come' have suppletive verb forms?	not included	
GB522	Can the S or A argument be omitted from a pragmatically unmarked clause when the referent is inferrable from context ("pro-drop" or "null anaphora")?	not included	

Table S2. WAIC values and quantiles (0.025, 0.5, and 0.975) of estimates of models predicting fusion and informativity (including the nonlinear implementations of L1 speakers and L2 speakers population): with random effects, with random and fixed effects, and with fixed effects.

model	response	effect	2.5%	50%	97.5%	WAIC
Phylogenetic+Spatial: local	fusion	phylogenetic SD	1.47	1.70	1.97	1,851.48
		spatial SD	0.29	0.35	0.42	1,851.48
		Intercept	-0.02	0.00	0.03	1,851.48
Phylogenetic		phylogenetic SD	2.01	2.22	2.45	1,956.18
		Intercept	-0.02	0.00	0.03	1,956.18
Phylogenetic+Areal		phylogenetic SD	1.82	2.04	2.30	2,032.46
		areal SD	0.08	0.15	0.27	2,032.46
		Intercept	-0.06	0.03	0.13	2,032.46
Phylogenetic+Spatial: regional		phylogenetic SD	1.25	1.49	1.79	2,086.63
		spatial SD	0.48	0.62	0.80	2,086.63
		Intercept	-0.02	0.00	0.03	2,086.63
Spatial: local		spatial SD	0.73	0.78	0.84	2,304.12
		Intercept	-0.02	0.00	0.03	2,304.12
Spatial: regional		spatial SD	0.90	1.07	1.25	2,456.76
		Intercept	-0.03	0.00	0.04	2,456.76
Areal		areal SD	0.44	0.58	0.79	2,909.68
		Intercept	0.08	0.31	0.54	2,909.68
Phylogenetic+Spatial: local+L1 speakers (linear)		phylogenetic SD	1.48	1.71	1.99	1,822.66
		spatial SD	0.29	0.35	0.42	1,822.66
		Intercept	-0.03	-0.01	0.02	1,822.66
		L1	0.04	0.09	0.14	1,822.66
Phylogenetic+Spatial: local+L1 speakers (nonlinear)		phylogenetic SD	1.48	1.71	1.99	1,823.02
		spatial SD	0.29	0.35	0.42	1,823.02
		Intercept	-0.02	0.00	0.03	1,823.02

		social SD: L1	0.00	0.01	0.03	1,823.02
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)		phylogenetic SD	1.49	1.72	1.99	1,823.56
		spatial SD	0.29	0.35	0.42	1,823.56
		Intercept	-0.17	-0.01	0.15	1,823.56
		social SD: L1	0.00	0.01	0.04	1,823.56
		social SD: L2 proportion	0.01	0.03	0.08	1,823.56
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)		phylogenetic SD	1.48	1.71	1.99	1,824.81
		spatial SD	0.29	0.35	0.42	1,824.81
		Intercept	-0.03	-0.01	0.02	1,824.81
		L1	0.04	0.09	0.14	1,824.81
		L2 proportion	-0.37	-0.03	0.31	1,824.81
Phylogenetic+Spatial: local+Neighbours		phylogenetic SD	1.47	1.70	1.97	1,849.87
		spatial SD	0.29	0.35	0.42	1,849.87
		Intercept	-0.02	0.00	0.02	1,849.87
		Neighbours	-0.01	0.02	0.05	1,849.87
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)		phylogenetic SD	1.46	1.69	1.96	1,853.72
		spatial SD	0.29	0.35	0.42	1,853.72
		Intercept	-0.02	0.00	0.02	1,853.72
		L1*L2 proportion	-0.03	0.04	0.11	1,853.72
Phylogenetic+Spatial: local+L2 proportion (nonlinear)		phylogenetic SD	1.47	1.69	1.97	1,854.31
		spatial SD	0.29	0.35	0.42	1,854.31
		Intercept	-0.16	0.00	0.16	1,854.31
		social SD: L2 proportion	0.00	0.01	0.04	1,854.31

Phylogenetic+Spatial: local+L2 proportion (linear)		phylogenetic SD	1.47	1.70	1.97	1,854.33
		spatial SD	0.29	0.35	0.42	1,854.33
		Intercept	-0.02	0.00	0.03	1,854.33
		L2 proportion	-0.34	0.00	0.34	1,854.33
Phylogenetic+Spatial: local+Official		phylogenetic SD	1.45	1.67	1.95	1,854.91
		spatial SD	0.29	0.36	0.42	1,854.91
		Intercept	-0.03	-0.01	0.02	1,854.91
		Official status	0.00	0.16	0.33	1,854.91
Phylogenetic+Spatial: local+Education		phylogenetic SD	1.45	1.68	1.95	1,855.62
		spatial SD	0.29	0.35	0.42	1,855.62
		Intercept	-0.03	-0.01	0.02	1,855.62
		Education	0.00	0.14	0.28	1,855.62
L1 speakers (nonlinear)		Intercept	0.09	0.17	0.25	3,654.92
		social SD: L1	0.07	0.17	0.55	3,654.92
L1 speakers (nonlinear)+L2 proportion (nonlinear)		Intercept	-0.03	0.26	0.54	3,656.95
		social SD: L1	0.05	0.14	0.75	3,656.95
		social SD: L2 proportion	0.01	0.03	0.08	3,656.95
Education		Intercept	-0.09	-0.03	0.02	3,676.87
		Education	0.25	0.45	0.65	3,676.87
Official		Intercept	-0.08	-0.02	0.03	3,680.30
		Official status	0.25	0.49	0.72	3,680.30
L1_log10:L2 proportion (linear)		Intercept	-0.07	-0.01	0.04	3,688.91
		L1*L2 proportion	0.06	0.18	0.29	3,688.91

L2 proportion (nonlinear)		Intercept	-0.04	0.24	0.53	3,692.84
		social SD: L2 proportion	0.00	0.01	0.03	3,692.84
L2 proportion (linear)		Intercept	-0.06	-0.01	0.05	3,694.48
		L2 proportion	-0.06	0.54	1.13	3,694.48
L1 speakers (linear)+L2 proportion (linear)		Intercept	-0.06	-0.01	0.05	3,694.91
		L1	-0.02	0.04	0.10	3,694.91
		L2 proportion	-0.11	0.49	1.09	3,694.91
L1 speakers (linear)		Intercept	-0.06	0.00	0.05	3,695.00
		L1	-0.02	0.05	0.11	3,695.00
Neighbours		Intercept	-0.05	0.01	0.06	3,697.24
		Neighbours	-0.06	-0.01	0.03	3,697.24
Phylogenetic+Spatial: local	informativity	phylogenetic SD	0.88	1.19	1.56	3,228.77
		spatial SD	0.36	0.44	0.54	3,228.77
		Intercept	-0.01	0.03	0.07	3,228.77
Phylogenetic		phylogenetic SD	1.47	1.77	2.11	3,277.79
		Intercept	-0.01	0.03	0.07	3,277.79
Phylogenetic+Spatial: regional		phylogenetic SD	0.63	0.95	1.37	3,298.40
		spatial SD	0.49	0.65	0.86	3,298.40
		Intercept	-0.01	0.03	0.07	3,298.40
Phylogenetic+Areal		phylogenetic SD	1.29	1.62	2.01	3,299.01
		areal SD	0.05	0.14	0.29	3,299.01
		Intercept	-0.05	0.04	0.15	3,299.01
Spatial: regional		spatial SD	0.56	0.72	0.90	3,381.21
		Intercept	-0.02	0.03	0.07	3,381.21
Spatial: local		spatial SD	0.57	0.64	0.72	3,381.62
		Intercept	-0.01	0.03	0.07	3,381.62

Areal		areal SD	0.30	0.42	0.59	3,511.16
		Intercept	-0.02	0.16	0.35	3,511.16
Phylogenetic+Spatial: local+L1 speakers (linear)		phylogenetic SD	0.97	1.28	1.65	3,207.14
		spatial SD	0.35	0.43	0.53	3,207.14
		Intercept	-0.03	0.01	0.05	3,207.14
		L1	0.05	0.13	0.20	3,207.14
Phylogenetic+Spatial: local+L1 speakers (nonlinear)		phylogenetic SD	0.97	1.29	1.65	3,207.49
		spatial SD	0.35	0.43	0.53	3,207.49
		Intercept	-0.01	0.03	0.07	3,207.49
		social SD: L1	0.00	0.01	0.03	3,207.49
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)		phylogenetic SD	0.98	1.30	1.65	3,208.11
		spatial SD	0.35	0.43	0.53	3,208.11
		Intercept	-0.24	0.01	0.26	3,208.11
		social SD: L1	0.00	0.01	0.03	3,208.11
		social SD: L2 proportion	0.00	0.01	0.03	3,208.11
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)		phylogenetic SD	0.97	1.28	1.65	3,209.11
		spatial SD	0.35	0.43	0.53	3,209.11
		Intercept	-0.03	0.01	0.05	3,209.11
		L1	0.05	0.13	0.20	3,209.11
		L2 proportion	-0.54	-0.02	0.50	3,209.11
Phylogenetic+Spatial: local+Neighbours		phylogenetic SD	0.88	1.20	1.57	3,229.76
		spatial SD	0.36	0.44	0.54	3,229.76
		Intercept	-0.01	0.03	0.07	3,229.76
		Neighbours	-0.04	0.01	0.06	3,229.76
Phylogenetic+Spatial: local+L2 proportion (linear)		phylogenetic SD	0.88	1.20	1.56	3,229.95
		spatial SD	0.36	0.44	0.54	3,229.95

		Intercept	-0.01	0.03	0.07	3,229.95
		L2 proportion	-0.49	0.03	0.55	3,229.95
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)		phylogenetic SD	0.88	1.19	1.56	3,230.07
		spatial SD	0.36	0.45	0.54	3,230.07
		Intercept	-0.01	0.03	0.07	3,230.07
		L1*L2 proportion	-0.08	0.03	0.13	3,230.07
Phylogenetic+Spatial: local+L2 proportion (nonlinear)		phylogenetic SD	0.88	1.19	1.56	3,230.77
		spatial SD	0.36	0.44	0.54	3,230.77
		Intercept	-0.22	0.03	0.28	3,230.77
		social SD: L2 proportion	0.00	0.01	0.03	3,230.77
Phylogenetic+Spatial: local+Education		phylogenetic SD	0.85	1.16	1.53	3,233.24
		spatial SD	0.36	0.44	0.54	3,233.24
		Intercept	-0.03	0.01	0.06	3,233.24
		Education	-0.02	0.19	0.39	3,233.24
Phylogenetic+Spatial: local+Official		phylogenetic SD	0.83	1.14	1.52	3,233.87
		spatial SD	0.37	0.45	0.54	3,233.87
		Intercept	-0.03	0.02	0.06	3,233.87
		Official status	-0.01	0.22	0.46	3,233.87
Official		Intercept	-0.05	0.01	0.06	3,713.84
		Official status	0.16	0.39	0.63	3,713.84
Education		Intercept	-0.05	0.00	0.06	3,714.79
		Education	0.12	0.32	0.52	3,714.79
Neighbours		Intercept	-0.02	0.03	0.09	3,718.70
		Neighbours	-0.10	-0.06	-0.01	3,718.70
		Intercept	-0.18	0.11	0.41	3,722.85

L1 speakers (nonlinear)+L2 proportion (nonlinear)		social SD: L1	0.03	0.04	0.05	3,722.85
		social SD: L2 proportion	0.05	0.05	0.06	3,722.85
L1_log10:L2 proportion (linear)		Intercept	-0.03	0.02	0.08	3,723.47
		L1*L2 proportion	-0.06	0.06	0.18	3,723.47
L1 speakers (nonlinear)		Intercept	-0.02	0.03	0.10	3,723.58
		social SD: L1	0.00	0.01	0.03	3,723.58
L2 proportion (nonlinear)		Intercept	-0.18	0.11	0.40	3,723.94
		social SD: L2 proportion	0.00	0.01	0.03	3,723.94
L1 speakers (linear)		Intercept	-0.03	0.03	0.08	3,724.21
		L1	-0.08	-0.01	0.05	3,724.21
L2 proportion (linear)		Intercept	-0.03	0.02	0.08	3,724.24
		L2 proportion	-0.42	0.19	0.79	3,724.24
L1 speakers (linear)+L2 proportion (linear)		Intercept	-0.03	0.03	0.08	3,726.04
		L1	-0.08	-0.02	0.05	3,726.04
		L2 proportion	-0.40	0.20	0.81	3,726.04

Table S3. Percentage of variance explained by phylogenetic and spatial random effects in top-ranking models of fusion and informativity incorporating random effects and the number of L1 speakers

response	sociodemographic predictor(s)	variance for the Gaussian observations in %	variance for phy_id in %	variance for sp_id in %	WAIC
fusion	L1 speakers	5	91	4	1,822.66
	L1 speakers + L2 proportion	5	92	4	1,824.81
informativity	L1 speakers	22	70	8	3,207.14
	L1 speakers + L2 proportion	21	71	8	3,209.11

Table S4. WAIC values of spatiophylogenetic models of fusion and informativity with fixed and random effects using different priors

model	WAIC	prior	response
Phylogenetic+Spatial: local+L1 speakers (linear)	1,822.66	0.1	fusion
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	1,823.02		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	1,823.56		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	1,824.81		
Phylogenetic+Spatial: local+Neighbours	1,849.87		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	1,853.72		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	1,854.31		
Phylogenetic+Spatial: local+L2 proportion (linear)	1,854.33		
Phylogenetic+Spatial: local+Official	1,854.91		
Phylogenetic+Spatial: local+Education	1,855.62		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	1,830.25	0.01	
Phylogenetic+Spatial: local+L1 speakers (linear)	1,832.20		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	1,833.31		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	1,834.44		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	1,857.65		
Phylogenetic+Spatial: local+Neighbours	1,857.88		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	1,862.63		
Phylogenetic+Spatial: local+Official	1,863.33		
Phylogenetic+Spatial: local+L2 proportion (linear)	1,864.04		
Phylogenetic+Spatial: local+Education	1,864.14		
Phylogenetic+Spatial: local+L1 speakers (linear)	1,817.08	0.5	
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	1,817.17		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	1,819.10		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	1,819.68		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	1,842.35		
Phylogenetic+Spatial: local+Neighbours	1,843.91		

Phylogenetic+Spatial: local+Official	1,848.30		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	1,848.38		
Phylogenetic+Spatial: local+Education	1,848.88		
Phylogenetic+Spatial: local+L2 proportion (linear)	1,850.11		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	1,811.03	0.99	
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	1,814.38		
Phylogenetic+Spatial: local+L1 speakers (linear)	1,814.79		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	1,816.96		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	1,840.19		
Phylogenetic+Spatial: local+Neighbours	1,841.26		
Phylogenetic+Spatial: local+Official	1,845.83		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	1,845.95		
Phylogenetic+Spatial: local+Education	1,846.80		
Phylogenetic+Spatial: local+L2 proportion (linear)	1,847.04		
Phylogenetic+Spatial: local+L1 speakers (linear)	3,207.14	0.1	informativity
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	3,207.49		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	3,208.11		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	3,209.11		
Phylogenetic+Spatial: local+Neighbours	3,229.76		
Phylogenetic+Spatial: local+L2 proportion (linear)	3,229.95		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	3,230.07		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	3,230.77		
Phylogenetic+Spatial: local+Education	3,233.24		
Phylogenetic+Spatial: local+Official	3,233.87		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	3,205.37	0.01	
Phylogenetic+Spatial: local+L1 speakers (linear)	3,214.27		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	3,216.08		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	3,217.42		

Phylogenetic+Spatial: local+Neighbours	3,236.55		
Phylogenetic+Spatial: local+L2 proportion (linear)	3,238.63		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	3,238.68		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	3,238.79		
Phylogenetic+Spatial: local+Education	3,241.52		
Phylogenetic+Spatial: local+Official	3,242.53		
Phylogenetic+Spatial: local+L1 speakers (linear)	3,201.63	0.5	
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	3,202.34		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	3,203.69		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	3,205.24		
Phylogenetic+Spatial: local+Neighbours	3,223.94		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	3,225.58		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	3,225.73		
Phylogenetic+Spatial: local+L2 proportion (linear)	3,225.88		
Phylogenetic+Spatial: local+Official	3,226.08		
Phylogenetic+Spatial: local+Education	3,227.25		
Phylogenetic+Spatial: local+L1 speakers (linear)	3,198.98	0.99	
Phylogenetic+Spatial: local+L1 speakers (nonlinear)	3,199.70		
Phylogenetic+Spatial: local+L1 speakers (linear)+L2 proportion (linear)	3,200.89		
Phylogenetic+Spatial: local+L1 speakers (nonlinear)+L2 proportion (nonlinear)	3,202.00		
Phylogenetic+Spatial: local+Neighbours	3,220.92		
Phylogenetic+Spatial: local+L1_log10:L2 proportion (linear)	3,222.78		
Phylogenetic+Spatial: local+L2 proportion (linear)	3,223.07		
Phylogenetic+Spatial: local+Education	3,224.39		
Phylogenetic+Spatial: local+L2 proportion (nonlinear)	3,224.41		
Phylogenetic+Spatial: local+Official	3,225.27		

S1 Spatial effects

We tested three versions of spatial parameters. Two were based on precision matrices that incorporated spatial parameters that assumed diffusion/borrowing of features 1) at shorter distances, such several hundreds of kilometers ("local" set of parameters) and 2) at larger distances, such as several thousands of kilometers ("regional" set of parameters) (see Figure S1). Additionally, we fit random effects of group membership in one of the AUTOTYP areas. When examining top-ranking models of fusion and informativity, we find that these invariably include the spatial effects fitting the assumptions of the "local" version. This indicates that complex features associated with fusion and informativity are likely to be shared by languages that are not located too far from each other.

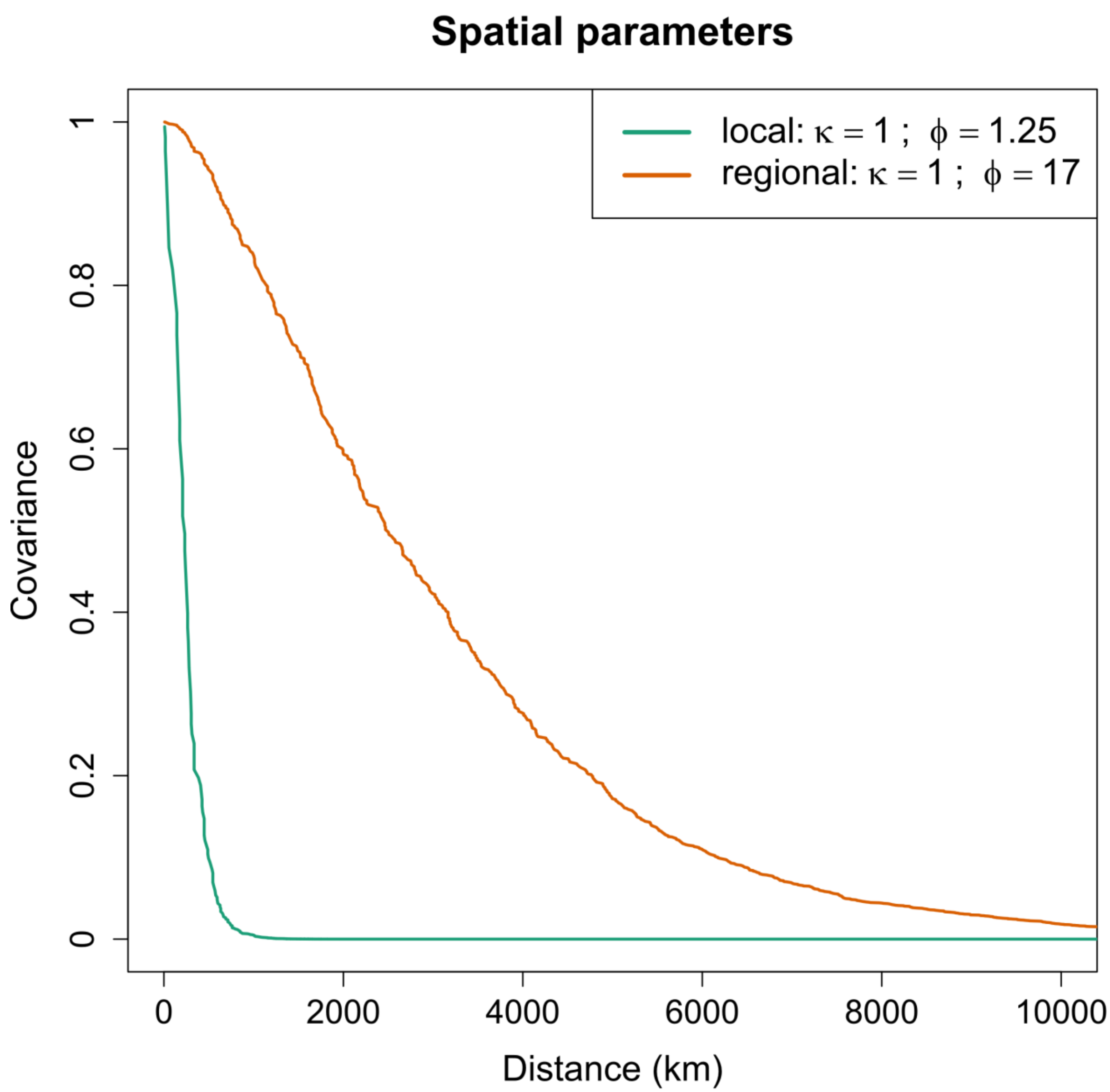


Fig. S1. The covariance in scores decays with increasing distances between languages: covariance decays faster between languages on shorter distances under the local set of spatial parameters (green) than on the regional set (red). The local parameters κ (the additional smoothness parameter) and ϕ (the correlation function parameter) restrict the diffusion between languages to distances below 1,000 kilometers, while under the regional parameters these distances can span several thousands of kilometers.