Year: 2017

Areas and universals

Bickel, Balthasar

Areas and universals*

Balthasar Bickel

1   Introduction

In explanations of how linguistic structures are distributed in the world, the pendulum has swung back from an emphasis on universals, that has dominated the second half of the twentieth century, to a renewed emphasis on local developments and areal diffusion. This shift in emphasis started over twenty years ago, with Dryer (1989), who drew attention to large-scale diffusion as an important possible confounding factor in the statistics of universals and with Nichols (1992), who set out to test universals but instead discovered an intriguing set of large-scale areal patterns. The general shift has gained further impetus with Maslova (2000), who raised the possibility that the current distribution of linguistic structures does not so much reflect universal tendencies but rather accidental skewings during the early developments of language, including skewings due to areal diffusion. Dediu & Levinson (2013) even suggest that the current distribution might reveal signals from ancient contacts with now-extinct sister lineages (Neanderthals, Denisovans and perhaps other unknown lineages) in Eurasia and the Sahul area.

The shift in emphasis has considerably challenged prospects for the discovery of universals. It has become clear that the sheer frequency of some patterns in the world does not in itself suggest effects from a universal principle of language: a pattern can be dominant in frequency worldwide just because it spread around the world through historically contingent cases of areal diffusion. For example, the fact that non-verb-initial languages are far more frequent than verb-initial languages does not necessarily suggest that the former are preferred as a matter of principle (e.g. because of a processing principle). The observed frequency distribution could just as well reflect a multitude of far-reaching but accidental spreads in the distant path (e.g. expansion of verb-final structure in the Eurasian steppe, spread of verb-medial order in Southeast Asia in the wake of Tai migrations, spread of verb-medial order together with the Bantu expansion etc.).

But how can we assess and demonstrate such kinds of areal diffusion? The problems here are by no means smaller than the challenges facing research on universals. First, the repeated spread of some pattern might itself be driven by universal forces: if a structure is universally preferred by some principle (e.g. by processing ease), we would in fact expect it to diffuse in contact situations more easily than universally dispreferred structures. For example, there is growing evidence that simple morphology, e.g. in case marking, is favored by adult learners (Bentz & Winter, 2013, summarizing earlier work), and there is tentative evidence from Artificial Language Learning experiments that certain word order patterns are more easily acquired by adults than others (Culbertson et al., 2012). Findings like these imply that some patterns spread more easily in the wake of second language

---

learning, i.e. in contact situations, than other patterns, possibly resulting in large-scale areal spreads.

A second problem is similar in kind to the possible fallacies when interpreting worldwide frequencies as universal preferences: if a pattern is more frequent in an area than outside it, this can be attributed to diffusion in contact only to the extent that we can be sure that the pattern did not arise many times independently because of some universal principle. For example, dependent marking on arguments (by case or adpositions) is known to occur with increased frequencies all-over Eurasia, with only few exceptions in Southeast Asia (Bickel & Nichols, 2009). This does not by itself establish that dependent marking spread in this area: it is equally possible that dependent marking is favored in the area independently in each family. This could be caused by the fact that dependent marking here is mostly embedded in verb-final sentence structures. These structures are expected to favor dependent marking universally (Greenberg, 1963; Nichols, 1992; Siewierska, 1996) because this kind of marking makes sentence processing (Hawkins, 2004) and/or meaning tranmission more efficient (Gibson et al. 2013; Hall et al. 2013).

Obviously, none of these problems can be solved by keeping the pendulum of emphasis active, sometimes attributing frequency patterns to areal diffusion, sometimes to universal trends. What is needed instead is an approach that focuses on the *interaction* between area formation and universals, seeking to identify the extents to which the two explain how linguistic structures are distributed in the world. This chapter discusses some possible ways of making progress in this question, drawing mostly on the theoretical framework of Distributional Typology (Bickel, 2015).

In the following, I begin the discussion by exploring the kinds of processes that lead to area formation and universal patterns (Section 2). This leads me to suggest ways of distinguishing the statistical footprint of the relevant processes (Section 3). Section 4 illustrates the methods via recent case studies, and Section 5 concludes the chapter.

2   Functional and event-based triggers of language change

As already noted, one of the key challenges in understanding areal patterns is that they need not result from historical contingencies but can just as well reflect universally preferred patterns of diffusion. The traditional opposition between areas and universals fails to capture the underlying processes and causes here.

As an alternative conceptual opposition, genetics offers the contrast between 'horizontal' and 'vertical' transfer, but these notions do not help much either: linguistic diffusion is always a vertical, diachronic process (Croft et al., 2011). When we say that a structure (say, verb-final order) has spread in an area, what is meant is that the languages in this area changed their structure so as to mirror the structure of their neighbors, or that they selectively kept structures that mirror those of their neighbors. Such processes can easily take several generations until their results stabilize in a community, and as Johanson (1992) has emphasized some time ago, these processes are fundamentally based on copying and imitating (calquing). Horizontal transfer in genetics, by contrast, involves the direct transfer of concrete

genetic material across synchronically co-existing individuals (as is common in bacteria).[1]

Given these conceptual problems, it is advisable to replace the traditional opposition between areas and universals by one that is grounded in the causes and conditions of distributional patterns. Approached from this perspective, the main contrast turns out to be between *functional* and *event-based* triggers (inducers) of language change.

Functional triggers are grounded in the biological/cognitive or social/communicative conditions of language, such as specific processing preferences (e.g. Hawkins, 2004; Christiansen & Chater, 2008) or specific sociolinguistic constellations (e.g. Trudgill, 2011; Lupyan & Dale, 2010) that systematically bias the way linguistic structures evolve. The defining property of functional triggers is that they affect transition probabilities universally, independent of concrete historical events. For example, if it is true that processing principles cause verb-final word order to associate with dependent marking, we expect this to cause a higher probability of languages changing towards than away from this association, and this transition probability is the same in any language, at any time.

But note that such kinds of universally fixed transition probabilities do not always result in universally wide-spread patterns: a functional trigger may be tied to a biological or social condition that has itself a limited distribution in the world. Possible such cases are discussed by Dediu & Ladd (2007) for genetic and Everett et al. (2015) for environmental bases of phonological tone, and by Evans (2003) for specific social structures as favoring specific grammars and lexicalizations in kinship expressions (known as 'kintax'). To the extent that the underlying functional trigger is indeed real, the probabilities for developing and maintaining the relevant structures (tone, kintax) are universally fixed even if they cannot yield the same results all over the world because the conditions do not obtain everywhere. When the conditions are met, however, functional triggers are expected to yield systematic, replicated patterns. Therefore, such triggers can also be called *principles*.

The expected systematicity of effects also entails that a functionally preferred structure can easily spread when languages are in contact, leading to area formation: if a certain structure, or a certain association between structures, is preferred by processing mechanisms or a certain sociolinguistic setting, there is every reason to expect that when speakers can select between variants in multilingual settings, they select the variant that complies with the preferred structure or the preferred association of structures.

Event-based triggers are tied to single historical events, leading to idiosyncratic, once-off changes. A classical example is the RUKI rule of Indo-European which retracted *s* after *r, *u, *k*, and *i* at some point in the history of the family. Event-based triggers are especially prominent in language contact situations, when a

---

1 Lexical borrowing would seem more similar to horizontal transfer than structural copying, but even lexical borrowing does not involve the direct transfer of Saussurian signs. Rather, the recipient language imitates a foreign word, using its own phonology, and so the process is more similar to the creation of a new word, inspired by a foreign model, but firmly under the constraints of the native sound inventory.

structure is copied not because it has a universally high probability of developing, but out of mere fashion. For example, relative pronouns and 'have'-based perfect tenses appear to have a very low probability of developing, as suggested by their extremely rare occurrence (cf. Comrie & Kuteva, 2013 and Dahl & Velupillai, 2013, respectively). As far as we know, the structures do not seem to be particularly preferred by either processing or social conditions of language use. However, where they occur, they appear to have spread in the wake of intensive language contact events, viz. in Europe during the transition period between antiquity and the middle ages (Haspelmath, 1998; Heine & Kuteva, 2006). Another example is the finding by Bickel et al. (2014) that differential object and differential subject marking are strongly dispreferred worldwide, but spread widely (to different extents) in two hotbeds, once in Southwestern Eurasia and once in Australia.

In these cases, event-based area formation rests mostly on the preferred development (grammaticalization) of a structure. But like in other diachronic processes, the key effect can also be the preferred maintenance, rather than innovation, of a structure. Examples of this are ergative case alignment or pronominal gender. These structures have been observed to rarely emerge, but when they do emerge, they tend to persist in areal clusters (cf. Nichols, 1993 and Bornkessel-Schlesewsky et al., 2008 on ergativity; Nichols, 1992 and Bickel, 2013 on gender).

What these examples suggest is that both functional and event-based triggers can lead to area formation. However, because functional triggers come with universally fixed transition probabilities, they can be expected to leave a different statistical footprint in the typological record than event-based triggers. Specifically, functional triggers can be expected to leave the same footprint across geographical regions. Event-based triggers, by contrast, are tied to specific historical contingencies, and so they are expected to leave signals only in single regions.

Quantification of this contrast is challenging because the relevant regions are often unknown. Since Dryer (1989) and Nichols (1992), typologists usually operate with a pre-defined set of regions (e.g. Africa, Eurasia, New-Guinea/Australia and the Americas, or more fine-grained distinctions). Functional triggers are then expected to leave the same statistical signal across all these regions while event-based triggers are expected to leave statistically different signals in each region. This is the basic state-of-the-art model in which we can test the effects of functional vs. event-based triggers, and I will discuss the relevant methods for this model and case study applications below. Before getting there, however, I wish to elaborate a bit further the challenges posed by the definition of regions.

Humans, linguists among them, are good at detecting spatial patterns. Browsing the *World Atlas of Language Structures* (WALS, Dryer & Haspelmath, 2013), for example, one can easily discern dozens of potentially interesting clusterings in space. What is not so easy is to tease apart genuine from spurious clusters. Statistical methods are only of limited help here because a statistical correlation can of course itself be spurious (as was recently highlighted for typological data by Roberts & Winters, 2013). What is needed is robust causal theories that motivate specific scenarios of event-based triggers and the distributional patterns that can be predicted from these scenarios, i.e. what are called *Predictive Areality Theories* by Bickel & Nichols (2006). The key idea of Predictive Areality Theories is that areal patterns are predicted by what is known from social/cultural history and

archeology, from language spreads and contact events, and from migration patterns as revealed for example through population genetics. An example is the Predictive Areality Theory of the Eurasian area which has good support in both the historical record of ancient migrations and language spreads and the population genetic record. Together, these records suggest multiple waves of intensive language contact events over at least the past 14'000 - 19'000 years (for a summary of the historical record, see Nichols, 1998, and for the genetic record, Rootsi et al., 2007).

3   Methods

The discussion so far suggests that the relative impact of functional and event-based triggers can be best evaluated through statistical modeling, following the well-established framework of *Generalized Linear Models* (cf. Baayen, 2008 and Johnson, 2008 for linguistically-oriented introductions): the distribution of some structure (say a specific word order) is predicted by the interaction of the conditions that are caused by the assumed triggers, for example, conditions of area formation that are caused by specific historical contact events (e.g. Eurasia) and conditions of universal preferences that are caused by functional triggers (e.g. verb-final clause structures favoring certain argument-marking patterns, or social organization types favoring certain 'kintax' patterns).

There are various ways of implementing such models, and there is no consensus yet which implementation is most suitable. The literature includes logistic (Sinnemäki, 2010) as well as log-linear (Poisson) models (Cysouw, 2010a; Bickel, 2011), and some authors model areas as random factors (Cysouw, 2010a; Jaeger et al., 2011), others as fixed factors (Bickel et al., 2009; Sinnemäki, 2010). In practice, these choices often have relatively little impact on results. Still, we clearly need more 'meta-typological' research of the kind presented in Maslova (2008) and Cysouw (2010b) to resolve the underlying issues.

What is far more pressing, however, is the question of how one can in fact assess whether a hypothesized diachronic process – functional or event-based, alone or through interaction – leaves a signal in the typological record. Synchronic frequency counts can be deceptive because we cannot assume that the distributions of our samples have reached what is known as stationarity (Maslova, 2000; Cysouw, 2011; Dediu & Cysouw, 2013). Stationarity means that when individual languages change types, this is compensated for by other changes elsewhere in the population of languages so that the total frequencies of types remain the same. If this is the case, synchronic frequencies provide a direct and reliable estimate of transition probabilities. But such a situation cannot be assumed *a priori* for typological data, especially since event-based triggers do not lead to universally fixed and constant transition probabilities (Maslova, 2000). Further, empirical studies on large families have shown that synchronic frequency distributions can indeed sometimes differ markedly from the frequencies that one would expect under stationarity (see for example Cysouw, 2011 for a demonstration of this with regard to word order and adposition patterns in Austronesian).

In response to these challenges, current approaches have moved away from interpreting raw synchronic frequency patterns and now aim at estimating trends in transition probabilities within individual families (from which in turn one can

calculate the expected stationary distribution for each family). There is a number of proposals that are currently being explored, but they all share the basic insight that differences and trends in transition probabilities can be estimated from languages that belong to the same family. If we know that some languages belong to the same family (because the family has been demonstrated using the Comparative Method), we know that the properties of these languages have developed through a chain of innovation and retention from a single proto-language (corresponding to what is known as a Markov chain in mathematics). The challenge then is to estimate the development of this chain. Some methods rely on pairs of languages that belong to the same family each, sampled at the same time depth (Maslova, 2004; Maslova & Nikitina, 2007; Dediu & Cysouw, 2013), others rely on exhaustive samples of languages in each family, at variable time depths (Bickel, 2011, 2013), and still others on estimated or reconstructed tree topologies, of variable size and time depth (Dunn et al., 2011; Dediu & Levinson, 2012).[2]

Each method has its advantages and disadvantages. Pair-wise methods make the problematic assumption that within an assumed time depth, there was at most one change and no reticulation (Dediu & Cysouw, 2013), but they have the advantage that one needs to sample only two languages per family. Exhaustive-sample methods require more data. This is a seeming disadvantage, but it has the advantage that the method can pick up more signals. Tree-based methods have the highest resolution, but they limit historical models to trees and exclude the possibility of wave and linkage models although such models are often plausible, especially from the point of view of areal diffusion processes (François 2014). Another requirement of tree-based methods is estimates of branch lengths, as these provide the relevant time intervals within which rates of change are calculated. This can be problematic when, as is usually done, branch length estimates are gained from lexical data (cognate replacements), for lexical change is not necessarily a good calibration stick for structural change. After all, lexical stems can be conservative while syntax may at the same time change rapidly, as one would precisely expect in fact for substrate effects (Thomason & Kaufman, 1988).

A key problem faced by all methods is that they require samples from within families. Yet about half of the known families on our planet are represented in available databases by only a single member (either because no other members are known at all or because no relevant data are available). For testing theories with universal scope or large-scale area effects, data from these single-member families can be absolutely critical. The only proposal so far that attempts to solve this problem is based on an extrapolation algorithm (Bickel, 2013).

The basic strategy is to first perform estimates on diachronic biases (differences in transition probabilities) within larger families, using one of the methods mentioned above. This results in an overall estimate on the probability for a family to show a diachronic bias (i.e. a significant difference in estimated transition probabilities) vs. being diverse (with no significant difference in the direction of

---

2 Software implementation is available and documented for all methods. Maslova's pair-wise method is available via Dediu & Cysouw (2013); my own exhaustive-sample-based `familybias` method via Bickel (2013). Tree-based methods rely on software developed for genetics, such as `BayesTraits` (Pagel & Meade, 2013), `MrBayes` (Ronquist & Huelsenbeck, 2003) or `geiger` (Harmon et al., 2008).

estimated transition probabilities). For the typological variable of Agent-Before-Patient vs. Patient-Before-Agent order, for example, there is a .69 probability for families to show a bias (in one of the two possible directions) and a .31 probability to be diverse (Bickel, 2013). Extrapolation then proceeds by randomly selecting 69% of the small families (with between one and, say, four languages) and declaring them as stemming from an unknown larger family with a bias. The actual value of the bias is read off the data (e.g. if Basque happens to be picked up as belonging to an unknown biased family, it would be taken to reflect a bias towards Agent-Before-Patient order, since this is the pattern shown by Basque). Allowing for statistical deviates (the sole known survivor of a family may be the odd one out, having undergone atypical development) and re-sampling the extrapolations in order to assess error margins, the algorithm then yields estimates on how many families are likely to have been biased vs. not biased diachronically, and when they are estimated to have been biased, how many families have been biased in which direction.

Another important challenge for all methods is that the statistical signals of diachronic biases is often very weak: functional principles and historical contact events can only ever trigger possible diachronies, but they do not themselves guarantee the emergence of actual developments: the actualization of a possible diachrony is entirely a matter of social propagation within a speech community. This process is fraught with multiple confounds of competing factors, ranging from competing functional principles and competing contact events to matters of social prestige and language ideology in a given community. As a result of this, it often takes many functionally-driven or event-based processes in individuals before the result of this can be detected in the diachronic biases of a language family.

4   Areal signals of Eurasia, hidden behind universals

A number of recent studies have re-assessed the evidence for functional principles that have been hypothesized to underlie statistical universals in the literature. While controlling for areal patterns, some of these studies have revealed evidence for event-based triggers of change that have affected distributions in addition to, but independently of, the relevant functional principles. These studies are interesting for the purposes of the present chapter because they suggest that evidence for event-based area formation can come from universals research. I focus here on two recent case studies that concern the Eurasian area.[3]

Performing log-linear analyses on diachronic biases, Bickel (2015) reports that families within Eurasia are significantly more likely to be biased towards developing and maintaining dependent marking on arguments (specifically, towards formal distinctions between agent and patient noun-phrases of two-argument clauses) than families outside Eurasia. This difference is independent of the equally

---

[3] See Sinnemäki (2010) for a case study showing evidence for Southeast Asia and for what Güldemann (2008) calls the 'Macro-Sudan Belt' in Africa, while testing a functional theory relating zero-marking and SVO order.

significant effect that leads verb-final families or sub-families[4] to biases towards dependent marking. In other words, the event-based diachrony that favors the emergence and maintenance of dependent marking in Eurasia is statistically independent of the functional principles that makes verb-final structures easier to process and/or transmit in communication if there is dependent marking than if there is no such marking (Hawkins, 2004; Hall et al., 2013; Gibson et al., 2013).

An earlier study (Bickel, 2011) looked at the Greenbergian correlation between the order of dependent noun phrases with respect to the head noun (left-branching [[NP] N] vs. right-branching [N [NP]]) and the order of patient noun phrases with respect to the verb (left-branching [[NP] V] vs. right-branching [V [NP]]), relying on Dryer's data in WALS (Dryer, 2005a, 2005b). The study controlled for various continent-sized areal patterns. However, since left-branching NPs are particularly frequent in the core of Eurasia (Masica's (2001) 'Indo-Turanian' area), it is possible that any areal effect on NP order is specifically caused by the difference between diachronic biases within vs. outside Eurasia. The Eurasian peripheries in Europe and Southeast Asia tend to show deviating NP orders (e.g. right-branching dependent NPs in French and Thai, respectively), but as argued in Bickel (2015) for other parts of grammar, deviating patterns like these may represent younger developments than those that shaped the Eurasian area as a whole. Incipient evidence for this possibility comes from the fact that left-branching NPs are attested even in the peripheries, sometimes as the dominant structure (Basque in Europe, Karen in Southeast Asia) and sometimes as an alternative pattern (English left-branching *John's house* vs. right-branching *the house of John*).

Based on these considerations, I estimated family biases in NP orders and performed a log-linear analysis of the association of these biases with VP order and the location within vs. outside Eurasia. The results are reported in Table 1.

Area	VP order	Biased families			Ratio of biases
		No bias	[[NP] N]	[N [NP]]	[[NP] N] / [N [NP]]
Inside Eurasia	[V [NP]]	3.67	7.93	14.40	.55
	[[NP] V]	6.86	45.30	0.84	53.93
Outside Eurasia	[V [NP]]	10.61	36.92	66.47	.56
	[[NP] V]	37.76	97.42	8.83	11.03

Table 1: Estimated diachronic biases of NP order in families with right-branching vs. left-branching VPs inside vs. outside Eurasia (data from Dryer 2005a, 2005b; estimation method from Bickel 2013 with extrapolation to isolates and small families).

The statistical analysis suggests that both area and VP order are significant and at the same time independent of each other.[5] This means that the family biases are

---

4 When families are split with regard to word order, as is the case for Indo-European or Sino-Tibetan, the study estimated biases within homogenous subgroups of the family. See the original study for details, and Bickel (2013) for theoretical justification.

5 This can be tested using a likelihood ratio test to compare models with vs. without the relevant interaction between variables: AREA × NP ORDER BIAS χ^2=4.96, *p*=.03; VP

best modeled by effects both from functional principles preferring harmony in branching direction and an event-based process that increases the number of families that are diachronically biased towards left-branching NPs in Eurasia. This can also be seen directly in the differences between the ratio of left-branching vs. right-branching families in the last column of Table 1. Inside Eurasia, families biased towards left-branching NPs outrank families biased towards right-branching NPs by a factor of almost 54. Outside Eurasia this factor is only about 11. No comparable difference can be observed with biases in right-branching NPs. This confirms the hypothesis that the Eurasian spreads specifically targeted left-branching NPs – not as a structure with a universal preference but as something that happened to be locally attractive for copying, a mere fashion.

The two case studies reported here show that critical evidence for event-based area formation can sometimes come from research on universal correlations. In both case studies, a hypothesized functional principle is confirmed by looking at estimates of diachronic biases in families. But the functional principles alone fail to explain the distribution sufficiently. In addition, there is evidence that event-based triggers of language change also significantly contributed to the observed spatial distribution.

5   Conclusions

Research on areas and research on universals have long been considered as irreconcilable opposites. The studies surveyed in this chapter have shown that both strands of research require each other. Research on areas is needed in order to control for alternative explanations when studying universals. Conversely, research on universals sometimes provides critical evidence for areal developments in particular regions. Progress in each domain requires a deeper and better understanding of the relevant trigger that causes the observed distributions: in-depth research on the cognitive/biological and social/communicative principles that favor the development and maintenance of specific structures, and at the same time, in-depth research on the historical and population-genetic events that intensified language contact during specific periods in specific regions.

Good causal theories in these domains make clear predictions on likely or less likely pathways of diachronic development. We have now a range of estimation techniques and databases at our disposal that make it possible to test these predictions. Good theories bring with them higher resolution in their predictions and modern estimation techniques require dense samples. This means that for all the methodological and theoretical progress that has been made on the question of areas and universals, the most urgent task remains the empirical groundwork of

---

ORDER × NP ORDER BIAS χ^2=112.2, *p*<.001. The three-way interaction is not significant (χ^2=1.79, *p*=.18), suggesting independence of the area and VP order effects. Family biases in NP order were estimated using the `familybias` function that is made available at `https://github.com/IVS-UZH/familybias` for use in `R` (R Development Core Team, 2013), with default parameter settings. When families are split in VP order, biases are estimated within non-split subgroups, just like in the other case study (see the previous note).

analyzing ever-larger arrays of languages across the world and making these analyses available in databases.

References

Baayen, R. Harald. 2008. *Analyzing Linguistic data*. Cambridge: Cambridge University Press.

Bentz, Christian & Bodo Winter. 2013. Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change* 3. 1–27.

Bickel, Balthasar. 2011. Statistical modeling of language universals. *Linguistic Typology* 15. 401–414.

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

Bickel, Balthasar. 2015. Distributional typology: statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine & Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis, 2nd edition*, 901–923. Oxford: Oxford University Press.

Bickel, Balthasar, Kristine Hildebrandt & René Schiering. 2009. The distribution of phonological word domains: a probabilistic typology. In Janet Grijzenhout & Bariş Kabak (eds.), *Phonological Domains: Universals and Deviations*, 47–75. Berlin: Mouton de Gruyter.

Bickel, Balthasar & Johanna Nichols. 2006. Oceania, the Pacific Rim, and the theory of linguistic areas. *Proc. Berkeley Linguistics Society* 32. 3 – 15.

Bickel, Balthasar & Johanna Nichols. 2009. The geography of case. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford Handbook of Case*, 479–493. Oxford: Oxford University Press.

Bickel, Balthasar, Alena Witzlack-Makarevich & Taras Zakharko. 2014. Typological evidence against universal effects of referential scales on case alignment. In Ina Bornkessel-Schlesewsky, Andrej Malchukov & Marc Richards (eds.), *Scales: a Cross-Disciplinary Perspective on Referential Hierarchies*, 7–43. Berlin: De Gruyter Mouton.

Bornkessel-Schlesewsky, Ina, Kamal Kumar Choudhary, Alena Witzlack-Makarevich & Balthasar Bickel. 2008. Bridging the gap between processing preferences and typological distributions: initial evidence from the online comprehension of control constructions in Hindi. In Andrej Malchukov & Marc Richards (eds.), *Scales (Linguistische ArbeitsBerichte 86)*, 397– 436. Leipzig: Institut für Linguistik [`http://www.uni-leipzig.de/~asw/lab/lab86/LAB86_Bornkessel_et_al.pdf`].

Christiansen, Morten H & Nick Chater. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31. 489–509.

Comrie, Bernard & Tania Kuteva. 2013. Relativization Strategies. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Leipzig: Max Planck Institute for Evolutionary Anthropology [`http://wals.info/chapter/s8`].

Croft, William, Tanmoy Bhattacharya, Dave Kleinschmidt, Eric D. Smith & T. Florian Jaeger. 2011. Greenbergian universals, diachrony and statistical analysis. *Linguistic Typology* 15. 433–453.

Culbertson, Jennifer, Paul Smolensky & Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122. 306–329.

Cysouw, Michael. 2010a. Dealing with diversity: towards an explanation of NP-internal word order frequencies. *Linguistic Typology* 14. 253–286.

Cysouw, Michael. 2010b. On the probability distribution of typological frequencies. In Christian Ebert, Gerhard Jäger & Jens Michaelis (eds.), *The Mathematics of Language*, 29–35. Berlin: Springer.

Cysouw, Michael. 2011. Understanding transition probabilities. *Linguistic Typology* 15. 415–431.

Dahl, Östen & Viveka Velupillai. 2013. The Perfect. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Leipzig: Max Planck Institute for Evolutionary Anthropology [`http://wals.info/chapter/68`].

Dediu, Dan & Michael Cysouw. 2013. Some structural aspects of language are more stable than others: a comparison of seven methods. *PLOS One* 8. e55009.

Dediu, Dan & Robert D. Ladd. 2007. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, *ASPM* and *Microcephalin*. *Proceedings of the National Academy of Sciences of the USA* 104. 10944–10949.

Dediu, Dan & Stephen C Levinson. 2012. Abstract profiles of structural stability point to universal tendencies, family-specific factors, and ancient connections between languages. *PLOS One* 7. e45198.

Dediu, Dan & Stephen C Levinson. 2013. On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology* 4:397. 1 – 17.

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

Dryer, Matthew S. 2005a. Order of genitive and noun. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 350–353. Oxford University Press.

Dryer, Matthew S. 2005b. Order of subject, object, and verb. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 330–341. Oxford University Press.

Dunn, Michael J., Simon J. Greenhill, Stephen C. Levinson & Russell D. Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473. 79–82.

Evans, Nicholas. 2003. Context, culture, and structuration in the languages of Australia. *Annual Review of Anthropology* 32. 13–40.

Everett, Caleb, Damián E. Blasi & Seán G. Roberts. 2015. Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences* 112. 1322–1327.

François, Alexandre. 2014. Trees, waves and linkages: models of language diversification. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge Handbook of Historical Linguistics*, 161–189. London: Routledge.

Gibson, Edward, Steven T. Piantadosi, Kimberly Brink, Leon Bergen, Eunice Lim & Rebecca Saxe. 2013. A noisy-channel account of crosslinguistic word-order variation. *Psychological Science* 24. 1079–1088.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, Mass.: MIT Press.

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

Hall, Matthew L., Rachel I. Mayberry & Victor S. Ferreira. 2013. Cognitive constraints on constituent order: evidence from elicited pantomime. *Cognition* 129. 1–17.

Harmon, Luke J., Jason T. Weir, Chad D. Brock, Richard E. Glor & Wendell Challenger. 2008. `geiger`: investigating evolutionary radiations. *Bioinformatics* 24. 129–131.

Haspelmath, Martin. 1998. How young is Standard Average European? *Language Sciences* 20. 271–287.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

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

Jaeger, T. Florian, Peter Graff, William Croft & Daniel Pontillo. 2011. Mixed effect models for genetic and aral dependencies in linguistic typology. *Linguistic Typology* 15. 281–320.

Johanson, Lars. 1992. *Strukturelle Faktoren in türkischen Sprachkontakten*. Stuttgart: Steiner.

Johnson, Keith. 2008. *Quantitative Methods in Linguistics*. London: Blackwell.

Lupyan, Gary & Rick Dale. 2010. Language structure is partly determined by social structure. *PLOS One* 5. e8559.

Masica, Colin. 2001. The definition and significance of linguistic areas: methods, pitfalls, and possibilities (with special reference to the validity of South Asia as a linguistic area). In Peri Bhaskararao & Karumuri Venkata Subbarao (eds.), *Tokyo Symposium on South Asian languages: contact, convergence, and typology [= The Yearbook of South Asian Languages and Linguistics 2001]*, 205–267. New Delhi: Sage Publications.

Maslova, Elena. 2000. A dynamic approach to the verification of distributional universals. *Linguistic Typology* 4. 307–333.

Maslova, Elena. 2004. Dinamika tipologičeskix raspredelenij i stabil'nost' jazykovyx tipov. *Vorprosy Jazykoznanija* 2004. 3–16.

Maslova, Elena. 2008. Meta-typological distributions. *Language Typology and Universals* 61. 199–207.

Maslova, Elena & Tatiana Nikitina. 2007. Stochastic universals and dynamics of cross-linguistic distributions: the case of alignment types. Ms. Stanford University, `http://anothersumma.net/Publications/Ergativity.pdf`.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: The University of Chicago Press.

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

Nichols, Johanna. 1998. The Eurasian spread zone and the Indo-European dispersal. In Roger Blench & Matthew Spriggs (eds.), *Archeology and language II: Archeological Data and Linguistic Hypotheses*, 220–266. London: Routledge.

Pagel, Mark & Andrew Meade. 2013. `BayesTraits` V2. Software and manual, `http://www.evolution.rdg.ac.uk/BayesTraitsV2Beta.html`.

R Development Core Team. 2014. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, `http://www.r-project.org`.

Roberts, Seán & James Winters. 2013. Linguistic Diversity and Traffic Accidents: Lessons from Statistical Studies of Cultural Traits. *PLOS One* 8. e70902.

Ronquist, Fredrik & John P Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19. 1572–1574.

Rootsi, S., L.A. Zhivotovsky, M. Baldovic, M. Kayser, I.A. Kutuev, R. Khusainova, M.A. Bermisheva, M. Gubina, S. Fedorova, A.M. Ilumäe, E.K. Khusnutdinova, L.P. Osipova, M. Stoneking, V. Ferak, J. Parik, T. Kivisild, P.A. Underhill & R. Villems. 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics* 15. 204–211.

Siewierska, Anna. 1996. Word order type and alignment. *Language Typology and Universals* 49. 149–176.

Sinnemäki, Kaius. 2010. Word order in zero-marking languages. *Studies in Language* 34. 869–912.

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

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.