



Erich R. Round* and Greville G. Corbett

Comparability and measurement in typological science: The bright future for linguistics

<https://doi.org/10.1515/lingty-2020-2060>

Abstract: Linguistics, and typology in particular, can have a bright future. We justify this optimism by discussing comparability from two angles. First, we take the opportunity presented by this special issue of *Linguistic Typology* to pause for a moment and make explicit some of the logical underpinnings of typological sciences, linguistics included, which we believe are worth reminding ourselves of. Second, we give a brief illustration of comparison, and particularly measurement, within modern typology.

Keywords: comparison, measurement, typology, substruction, linguistic description, Canonical Typology, Multivariate Typology, genitive, case, Russian

1 What typological science is like

Linguistic typology is part of a wider intellectual undertaking, in which we can benefit from the successes of others. In the early twentieth century, ‘typology’ was understood across the social and biological sciences as a family of scientific methods, which had emerged in the preceding decades and could be contrasted with ‘classification’ (Greenberg 1974 :11–17; Rudner 1966: 32–40). Classification, based on ‘classes’, was used in domains of inquiry, typically the hard sciences, where the categories, units, and scales of measurement to be used had by that time gained widespread acceptance and confidence, as when one refers unhesitatingly to carbon and oxygen, two-body and three-body gravitational systems, or mass and length. Typology, based on ‘types’, was for domains where

The authors contributed equally to the paper.

***Corresponding author: Erich R. Round [ɛɪɪk ɹaʊnd]**, University of Queensland, Brisbane, Australia; Max Planck Institute for the Science of Human History, Jena, Germany; and Surrey Morphology Group, University of Surrey, Guildford, England, E-mail: e.round@uq.edu.au
Greville G. Corbett [gɹɛvɪl kɔːbɪt], Surrey Morphology Group, University of Surrey, Guildford, England, E-mail: g.corbett@surrey.ac.uk

the categories, units and scales needed for the description and comparison of the ultimate object of inquiry were still under contention, and therefore additional methods were needed in order to build, discover or assess them (Hempel and Oppenheim 1936; Rudner 1966: 4–7; Weber [1904]1949).^{1,2} At its core, then, the typological method was not inherently about assigning objects to discrete categories (though this was an often-used technique), but about grappling in a scientific way with possible dimensions of comparison to use in the study of some domain of interest: how to discover them, evaluate them, and understand their contribution to the ultimate scientific goal of understanding the domain of inquiry. Owing to this history of ideas, since the late 1800s, an immense amount of philosophical and practical work has been carried out – within psychology, anthropology, sociology, archaeology and biology – into methodological questions which inherently still lie at the heart of linguistic typology today. Below we draw upon insights gained in other disciplines, when they apply equally in linguistics.

1.1 Domains, dimensions and debate: Typological sciences use them all

In practice, linguists study language topic by topic, domain by domain. But how do we define any given domain that we are studying? Unlike in a classificatory field, where it is a settled matter what it means to study carbon isotopes, in linguistics as in any typological field, merely justifying a delineation of a domain of study is itself a highly non-trivial task. Heated debates may persist for decades over the merits of competing delineations. As practising scientists, we often find ourselves needing to refine our own demarcation of the outer limits of a given domain as we better understand its internal make-up. Consequently, in practice we typically operate with ‘working definitions’, subject to revision (Cartwright and Bradburn 2011: 57). This is natural in a typological field. While modern-day chemists may not spill gallons of ink over the ‘best’ definition of carbon, linguists need to establish the

¹ ‘Classification’ is a term with many senses, including within linguistics. Here we use it strictly in this rather particular, early twentieth-century sense to help us frame some useful notions about scientific methodologies and what they are designed to handle. It is not our intention to comment on other things that have also been termed ‘classification’.

² Examples of ‘types’ in other fields include personality types in psychology (Gerlach et al. 2018; Jung 1921), categories of religions in anthropology (Johnson and Grim 2013; Murdock 1967), metrics such as the Human Development Index in economics (Stanton 2007) and the species concept in biology (de Queiroz 2005).

merits of possible categorizations. Debates about proposals (which, among other guises, can take the form of debates about terminology) are therefore inherently necessary for progress;³ debate and differences of view *per se* are not signs of confusion or disarray in a typological discipline. However, it is also true that there are more productive and less productive debates, and that the proper expectation even in typological sciences is to move towards greater agreement and certainty in the long run (Nettle 2018: 231–245). In this paper we highlight methods that we and others before us have advocated for increasing the productivity of the debates that typological disciplines have to have. In preparation we offer a quick note on the task of operationalization, and on the general notion of bringing coherence to a domain.

Supposing that as typologists we have in hand a working definition of a domain, an early task may be to operationalize it, that is, to state procedures which enable actual phenomena to be divided into those that lie inside and outside the domain. For example, suppose that in a typological study we were to choose to use Haspelmath's (2012: 123) definition of 'root' as "a morph that denotes a thing, an action or a property". For the sake of the argument, suppose also that we can distinguish 'morphs' unproblematically, and that ultimately every morph truly either does or does not denote a thing, action or property, such that given sufficient knowledge of any language there are no true grey areas. To carry out such a study, we would still require procedures for deciding about individual morphs in practice. Do we use a dictionary? Or a corpus? And selected according to what criteria? When available information is potentially incomplete, do we discard data or do we categorize all of it? If the latter, do we record the level of uncertainty of our decisions? These are matters of operationalization. In general, multiple operationalizations are possible and one can compare and contrast how faithfully they map back onto the original definition, a quality known as 'validity'. Some disciplines now have highly mature debates around definitions, operationalizations and the distinctions between them. See Feest (2005) for the philosophical debate in psychology over the last century, and Hull (1968) for a classic essay with respect to the species concept in biology. An accessible, practical overview of operationalization and validity is Ember et al. (1991).

Having selected a domain to study, such as nominalization, suppletion, or complex codas, we must determine which phenomena lie within the domain versus outside of it. Our next core task is to relate the phenomena that lie within the domain to one another, and to relate the variation among them to the overall

³ See Lakatos (1976) for a classic discussion of how progress on definitions (in mathematics, as it happens) is made through debate and fn. 25 regarding a contrary position proposed recently within linguistics.

possibilities that the domain notionally affords. We will want to ask: within the overall domain, do our observations fill the available space of possibilities completely, or do they leave pockets unoccupied? Are non-empty regions filled equally densely or are some sparser than others? To answer these questions well, we must be capable of coherently locating all the sub-parts of our domain, which in turn requires a concept of all the directions in which the domain extends, and how far. For example, what are the distinct dimensions of variation that exist in the domain? How is each of them divided up, and thus how is the domain as a whole ‘parceled out’ into smaller regions? Do we construe these dimensions as continuous, or divide them into discrete categories? If a dimension consists of discrete categories, how many categories are there, what are the boundaries, and are the categories unordered or ordered? If ordered, is the ordering linear or hierarchical, exhaustive or only partial? If a dimension is continuous, are there inherent limits at one or both ends, or is its extent (at least notionally) unlimited? Is there a choice among the scales that we could impose on it, such as linear, logarithmic or others? Quantitatively speaking, choices about how we divide up our dimensions will influence the kind of statistics available for analysing results (Stevens 1951 is a classic discussion). Turning from the internal division of dimensions, to questions around the selection of dimensions themselves, for each dimension that we are contemplating using, are there alternatives that do almost the same job – like magnetic north versus true north? If so, what are the ‘edge cases’ where the difference matters? What makes typology distinctively challenging, is that these very basic questions about the domains we wish to study often do not yet have settled answers; the field is still figuring them out. Consequently, we typologists are tasked with inventing our tape measures, and even defining the very dimensions in which we are to measure, at the same time as carrying out the empirical measurements of our objects of inquiry.

The notion that we should be engaged in inventing our tape measures at the same time as measuring our objects of inquiry can be disconcerting. It is worth recognizing first that this uneasiness is indeed real, then acknowledging why it comes about. If we accept that it is not only real but also valid, we can then consider what the appropriate response is. This ‘typological discomfort’ is a rational reaction to the fact that changes to tape measures (i.e. to our answers to the questions arrayed in the previous paragraph) will alter the implied, empirical relationships among our observations. As we change our answers, old clusters of ‘similar’ observations may split apart and new ones coalesce. What counts as a ‘true generalization’ under one construal may prove false under another. That ought to provoke discomfort. And typological discomfort is thus a valid reaction to the nature of our work. But in the face of it, we must hold our nerve. A faulty response to typological discomfort would be to capitulate, to believe that we cannot tolerate this burden any longer, and thus to

settle immediately on an answer, any answer, calling for all domains, dimensions, categories and scales to be frozen in place. Such a response, however, will not solve the essential problems that generate the abundance of choice that is provoking our unease; rather at best, it seeks to remove them from immediate view. Instead, a scientific, typological response is to seek out productive and critical debates about the options that currently appear to be open to us, to enable us to gradually discover which options generate the greatest insight and appear to map most securely onto the diversity of the world's languages.

In the remainder of the paper we discuss elements in what we regard as a toolkit for enumerating and evaluating the options that present themselves to us in typological research. As a prelude to that discussion, we have emphasized that much core scientific work in linguistics will consist in proposing and evaluating multiple conceivable ways of dividing up dimensions of linguistic variation, in identifying the dimensions themselves and considering their many alternatives, and in reevaluating the outer limits of domains.⁴

1.2 The productive comparison of scientific proposals

Linguistic studies do not unfold *ex nihilo*. Consequently, and especially given the typological nature of the field, progress will be made when we are best able to lucidly compare the results of competing typological studies, so as to derive evaluations of them that can lead to further improvement. Paul F. Lazarsfeld's (1937) *Some Remarks on the Typological Procedures in Social Research* was a seminal paper on the typological methodology of doing just this, and we begin with the central insights it contained.

Beyond describing and bringing order to what has already been observed, a typology – a scientific parcelling out of some domain space – should serve us by accomplishing two things. It should enable new observations to be placed coherently within it; and it should help us understand not merely the set of extant data points, but also the broader domain in which those data exist. As a method for ensuring that a typology, *T*, meets these desiderata, Lazarsfeld (1937) proposed a process he termed substruction. Substruction entails taking a typology, a parcelling of a domain into certain dimensions, and relating its dimensions to the underlying attributes of the domain. For example, we might

⁴ A complementary endeavour will be synthesizing work, that pursues dimensions which seem to recur across multiple domains (a classic case would be the Animacy Hierarchy, e.g. Corbett 2000: 54–88). These recurring dimensions help us relate domains themselves to one another, and suggest paths to potential integration and unification.

have typologized the inherent sonority of phonetic segments in terms of their overall spectral energy (Parker 2008). But while overall spectral energy is the dimension we have chosen, the space of phonetic segments has many more attributes: articulatory, perceptual, other acoustic attributes, including attributes that are more specific such as spectral energy in the range of 500–5,000 Hz. We may then have divided segments into four classes, ordered by energy level, whereas the nature of the underlying attribute (spectral energy) would also have allowed for the use of other discrete divisions, or a continuous scale, though it would not have lent itself, for example, to the use of four unordered categories. Substruction involves declaring that in parcelling out a domain into a series of dimensions, a typology T has utilized certain possible attributes and not others, and has divided them in some fashion, but not others. For the mundane progress of science, this task can be deeply useful, as it explicitly illuminates one of the central sets of reasons why two typologies T and U may differ, even in cases where the demarcation of their domain and their observational dataset are identical. And, by requiring that dimensions be related to attributes, substruction not only documents the fact that certain methodological decisions were made, but assists in articulating their significance (Lazarsfeld 1937: 133). For example, if three typologies of sonority produce different results, we would like to know the significance of the difference. Substruction may clarify that the first two differ only in setting different boundaries for certain acoustic measurements whereas the third also refers to articulation, thus the first two typologies appear merely to be operationalized differently, while typology number three is likely to be examining a different conceptualization of sonority.

Substruction also involves taking the parcelling out of a domain attribute by typology T, and relating that to the full, conceivable extent of the attribute. It asks, do the parcels properly exhaust the conceivable variation in the attribute? Do the parcels overlap? If the parcelling out of the space by typology T is non-exhaustive, then even if all points in the current dataset are covered, T may find new data unassignable. Conversely, if categories in T are overlapping, then even if all points in the current dataset have been uniquely assigned by it, new data may fall into a contradictory zone and be ambiguous (Lazarsfeld 1937:133). For example, in a phonetic typology, suppose we have posited a ‘sonorant’ category, whose members have prominent periodic harmonic energy, and a ‘fricative’ category, whose members have prominent higher-frequency aperiodic energy. This typology may work flawlessly for the first several hundred languages. However, critical attention to the underlying acoustic attributes would indicate that it should also be possible for a segment to have both kinds of energy, and indeed there are languages, such as Central Lisu (lis, Tabain et al. 2019) with

contrastive fricative vowels (Ladefoged and Maddieson 1990). Exactly how we should adjust the typology to categorize them is an open question; but it is the kind of question that substruction leads us to notice well before we happen upon a language that defies our categories. For typologists currently crafting a typology, substruction can thus offer immediate practical assistance by aiding the identification of such problems pre-emptively.⁵ In scientific debates on the relative merits of typologies T and U, it can help clarify the substance of the differences between the typologies, and hence why they give different results and what those differences tell us. As we remarked above, such scientific debates can appear in the guise of debates over definitions, terminology or interpretation, but essentially they are about typologies. Since these debates make an essential contribution to the advancement of typological science, we gain no long-term benefit by avoiding them or shutting them off prematurely, but rather when they are accompanied by meticulous substruction, we may reasonably expect them to unfold and conclude with more clarity, and more rapidly and fruitfully.

Substruction reveals insights not only about typologies, but about domains themselves. By demanding that a typology T be related to the underlying attributes of a domain, the substructive process requires us to state and justify, to ourselves and scientific peers, our beliefs about what those attributes are and how far they

5 A referee asks how it is possible to know pre-emptively whether some typology avoids these problems. We answer that substruction can *aid* us to do so but naturally, however hard we try, the linguistic data may prove yet more interesting than we imagined. If it is, our method can cope with that. Indeed, if we have used substruction in this way, then it tells us something about how remarkable the new data is, if it does not fit into our typology. For example, Corbett (2015) establishes a typology of inflectional phenomena, in terms of the ways in which lexemes can be split. Four criteria were justified, each with two possibilities; since the four criteria are orthogonal, there are in principle 16 possible combinations. The 16-member typology proved remarkably complete, with evidence for all possibilities occurring. Furthermore, it was suggested that the four criteria did not just sample the theoretical space but exhausted it, and arguments were given that could be taken as a substruction (2015: 177–178). The typology seemed secure, being complete, with the logic of its construction spelled out. A version was presented in Canberra, in November 2016. A year later, having heard that account, Don Daniels suggested that he had found data in the tense system of the Papuan language Soq (mdc), which seemed not to fit. This was not a simple gap in the combinatoric possibilities of the criteria previously established, but an entirely new dimension of variation. After careful analysis of ‘repartitioning’, as evidenced by Soq tense, a new, extended typology was proposed (Daniels and Corbett 2019: 734–742). This typology has several empty cells, raising again the issue of what is possible. In sum, the original typology seemed secure, and though substruction was used to check for gaps, the Soq phenomenon did not fit, and is indeed remarkable; prior to observing it we would not have considered it possible. We cannot be entirely sure pre-emptively that our typologies will be problem-free, but by using a substructive approach, our typologies should be well suited to bringing such phenomena to light.

conceivably extend. Doing so may reveal, for example, that two typologists agree on the underlying attributes of a domain, but have built the dimensions of their typologies differently (as in our first two sonority typologies); or it may reveal a disagreement over the assumptions about the attributes themselves (as in the example, above, of sonority typology number three versus the other two). Explicit statements about underlying attributes can also generate useful empirical hypotheses about data points that may exist but have not yet been observed. Lacunae in our current observations may become obvious once we clearly characterize the attribute space of the domain (Lazarsfeld 1937: 133–136). Characterizing the full conceivable extent of an attribute may suggest that even more extreme data points could exist beyond what has been observed so far. The upshot is that even in the absence of an explanatory theory of why we have observed what we have, substruction enables us to clarify our beliefs about the potential empirical context of those observations, and it reveals testable empirical hypotheses about what is yet to be observed.

1.3 The conceivable, the possible, and explanatory linguistic theories

At several points we have mentioned the comparison of typologies and observations on the one hand, with the conceivable extent of an attribute on the other. A distinction can be made between the conceivable and the possible.⁶ Take, for example, a language in which every nominal lexeme has thousands of inflected word forms, every one of which is a bare, suppletive stem; such a language is conceivable, but probably not possible. A language which has a full expressive capacity, but all of whose lexemes have just one, shared form /ta/, is not conceivable. The question of what is conceivable is logically prior to the question of what is actually possible. And frequently, while it is relatively simple to delineate what is conceivable, the question of what is possible may be one whose answer is unknown; it is likely to be determined by any number of complicating factors, many of which are poorly understood. Nevertheless, possibility is something of keen interest: after all, the ultimate question of linguistic science is not ‘what is a conceivable language?’, but ‘what is a possible language, and why?’ The path to an explanatory theory of possible languages runs in large part along the empty gap between what is conceivable and what is observed. By considering what is conceivable yet unattested, we are prompted to formulate testable theories that explain the gap. Thus we emphasize the important role of the

⁶ For an extended and instructive discussion see Dennett (1996: 103–107); for a generative perspective on possible and probable languages, see Newmeyer (2005).

conceivable, and the contribution to scientific knowledge that comes from considering domain attributes not only in terms of observed traits, but also their conceivable extents. More generally, we have attempted to convey the manner in which substruction facilitates the path from questions of ‘what are the dimensions of variation?’, to ‘what is the conceivable variation?’, to ‘what is possible, and why’, while concomitantly providing tools for comparing and evaluating the variety of answers that our discipline collectively proposes – a variety which a typological field such as linguistics inherently needs.

1.4 Grammar writing as a typological undertaking

Analysing the grammar of a single language, just like doing cross-linguistic typology, demands a typological method, in the sense we introduced at the start of Section 1, not a classificational one. Dimensions of variation are generally not self-evident but need to be identified, while a selection needs to be made among competing possible characterizations. Unlike cross-linguistic typology, the database for a single grammar is dominated by just one language, though in reality, considerations from nearby and related languages frequently play a role, as will broader cross-linguistic generalizations (see Evans and Dench 2006). The result is that language-particular categories – the parcelling-out of the space of variation in one language’s forms and functions – can be highly language-specific. At the same time, the attributes that underlie that space are often shared across languages. Consequently, the variation found in the grammar of one language can often be related meaningfully to cross-linguistic typologies, especially when the comparison is framed in relation to underlying attributes, of the kind that are identified during substruction.

1.5 Comparison

As in any science, one of the greatest sources of knowledge in linguistics is comparison. Through comparisons we find similarities and differences, and patterns among them that suggest explanatory forces that shape language. Earlier we mentioned the dimensions of typologies and attributes of a domain. A first step toward discovering possible dimensions and attributes is to compare two objects in a domain and ask in what manners they differ. Each manner, or axis, of variation may suggest a dimension or attribute. To find patterns, we compare comparisons. Doing so is aided by having a set of typological dimensions at hand. However, if our research is still at the early point of teasing out typological dimensions, then a useful stand-in is a point of reference. Points of reference are

tools of scientific inquiry used in both typological and classificational fields. They enable all other objects in a domain to be compared back to the same origin. Gaining an overview of how all those objects compare against the point of reference is a specially powerful resource in typological fields when the task underway still involves sorting out dimensions and attributes. Consequently, it makes sense to ask what kinds of points of reference are available, and what their affordances and limitations are. Here we briefly distinguish between three classic kinds of point of reference: standards, sample summaries, and idealizations. We then add one additional concept, a canon.

Standards are real-world properties or objects within a domain (Lazarsfeld 1937: 122–126). For example, one hundred degrees Celsius is a standard, defined as the temperature at which water boils; and for over a century, one metre was the length of a specific bar of platinum housed in the Archives nationales in Paris. Other properties and objects can then be compared to these real-world standards. Standards are useful for the task of comparison, but they have drawbacks. Two main drawbacks of standards arise due to the inevitable complexity of the real world from which any standard is drawn. First, a real-world standard may vary according to its context. One hundred Celsius is defined not only as the temperature at which water boils, but specifically water at an atmospheric pressure of one Standard Atmosphere and with the isotopic composition of Vienna Standard Mean Ocean Water. Due to the potential for real-world contexts to vary, the successful definition of one standard may rely on its context having previously been pinned down by the successful definition of yet other standards, making the initial establishment of standards potentially difficult. Second, real-world standards will have very many properties which are irrelevant to their status as a standard. A piece of string is compared to the platinum bar in the Archives nationales not in terms of its colour, stiffness, cost, smell or weight, but in terms of its length. Length, though, is a well-recognized and understood attribute, so at least we can be fairly certain what characteristic we want to compare. In typological fields, it can be much harder to prescribe precisely which traits of a real-world standard are intended to be relevant.⁷ Standards, therefore, come with the attendant ambiguities and difficulties of ostensive definitions.

Sample summaries are abstractions over many real-world objects. They include averages, but also more complex notions like prototypes and stereotypes.

⁷ For example, a classic case of measurement based on standards is Moh's scale of mineral hardness, which employs no fewer than 10 standards, such as talc, quartz and diamond. To place a mineral on the scale, one scratches it against the standard minerals, and the harder one of the two scratches the softer. This is feasible for minerals in large part because it is uncontentious what 'talc', 'quartz' and 'diamond' are; they are not changing over time; and both the performance and interpretation of the scratch test are easy and clear. In most typological domains however, such ease and clarity are lacking, and for that reason the use of standards is more fraught.

While everyday human cognition may make significant use of prototypes and stereotypes (Levinson 2000; Rosch 1978), their utility as scientific points of reference is hampered by their instability.⁸ The properties of sample summaries derive from properties of a particular empirical sample. If the empirical sample changes, for example, as we acquire a more complete understanding of linguistic diversity, then the summaries will change also. Naturally, one might freeze the sample, in effect creating a standard out of it, but this returns us to the problems of standards.

Idealizations have a long history of utility in science, from the frictionless plane, to absolute zero temperature, to the rational economic actor. The utility of idealizations in typological fields was first emphasized in detail a century ago by Max Weber ([1904] 1949). Idealizations avoid the messiness of the real world, while retaining the ability to be compared to real-world objects. One type of idealization is the logical extreme. A commonly encountered logical extreme is the complete absence of a property, such as friction in physics, or voicing in linguistics. Whether or not such an absence is actually instantiated in any real-world objects is immaterial; the utility of logical extremes is that they mark a fixed, hard limit to the conceivable variation in some attribute. This is not to say that idealizations are easily come by, or that all imaginable idealizations are equally useful (Rudner 1966: 54–63). The extreme idealization of absolute zero received acceptance only after centuries of related progress in thermodynamics (Wisniak 2005). And while a hyperspace drive is also an idealization, it is an idealization whose utility to science is unclear at best. Useful idealizations need not be extremes, but can also build on the notion of exactly one unit, such as the charge of exactly one electron, a verb with exactly one argument, or a morphosyntactic property with exactly one phonological exponent (see further in Section 2.3). These notions may in fact be instantiated in the real world, but they are not standards: their definition does not hinge upon which specific electron, or verb or morphosyntactic property is being referred to, but upon the fact that there is precisely one of them. It would be mistaken to claim that idealizations are a

8 A reviewer comments that surely ‘the most useful aspect of picking out a point of reference is to measure typicality’, which raises the crucial distinction between the establishment of a measure and the subsequent use of it. If we were establishing a measure to describe the average height of adult women, it would be a poor choice to set the ‘zero’ reference point as equal to the average itself and then measure individuals relative to that. Why? Because ‘zero’ would then move as the population itself changes (‘zero’ in 1960 would not equal ‘zero’ today), or even as our information about the same population changed. Instead, a general measure of height is established relative to the logical extreme point of 0 cm. Having established it there, we can then use it to characterize the height of individuals, and if we choose, to summarize those measurements into statements about typicality (see also fn.23). (One might counter, that IQ is defined in terms of the population’s mean performance on a certain test, and so is exactly like our example of mean women’s height, but in measurement theory IQ is criticised for having precisely that property: “[t]his form of characterization makes knowledge accumulation difficult, since by definition no other procedure can be used to measure the same quantity” (Cartwright and Bradburn 2011: 56)).

panacea come to cure all of the difficulties of comparison, or that finding good idealizations is without its challenges. For example, the extent to which a unit-idealization is useful may depend on the clarity of the definition of the unit. However, idealizations bring with them certain inherent advantages over standards and sample summaries.

In our discussion of points of reference, we have now considered standards, summaries and idealizations. We conclude this section of the paper with the additional notion of a canon.

Canons are composites of multiple points of reference. For instance, the canon for an agreement system would make reference to agreement controllers, targets, and so on. Because a single canon contains multiple points of reference, each associated with an axis of variation, it enables us to compare objects within our domain along multiple dimensions at once. In Section 2 we provide an example of how canons can be used in practice. One might wonder why canons are necessary, over and above the individual reference points they subsume. The reason is that in linguistics most of the objects of interest to us are multidimensional. They vary along multiple axes, all of which we would ideally like to measure. Canons help us do that by gathering together a coherent set of dimensions and reference points that are relevant for a domain of interest. By design, canons are not created to replace existing concepts in linguistic typology. Instead, they enable one to relate existing knowledge, or disputes,⁹ to a set of dimensions that facilitates unambiguous empirical comparison; canons thus promote the kind of productive debate over important typological matters which substruction enables. Accordingly, a canon will typically correspond to a set of dimensions that the field has discussed previously, while standing ready for revision and expansion as required (Evans et al. 2018a, 2018b, especially 2018b). Canons therefore enable a quick, coherent entry into the kind of rigorous typologizing advocated above, and they facilitate productive engagement with existing debates in the field. Typological domains which have been successfully examined in this way include agreement, negation, quotation, reality status, finiteness, tone, stress and phonaesthemes.¹⁰ Canons

⁹ As Nikolaeva (2013: 100) emphasizes, '[t]he canonical approach breaks down complex concepts in a way that clarifies where disagreements may lie between different linguists and theoretical frameworks'; see also Kwon and Round (2015).

¹⁰ Sources include, just as examples: agreement (Corbett 2006), tone and stress (Hyman 2012), negation (Bond 2013), quotation (Evans 2013), finiteness (Nikolaeva 2013), reality status (Michael 2014), inflection (Corbett 2015), phonaesthemes (Kwon and Round 2015), gender (Corbett and Fedden 2016), morphological complexity (Stump 2017), concurrent feature systems (Round and Corbett 2017; Fedden and Corbett 2017), overabundance (Thornton 2019), compounding (Spencer 2019). Canons prove useful in the analysis of signed languages too (Cormier et al. 2013). For a recent overview, concentrating on morphology, see Bond (2019), and for further examples of the value of canons see the bibliography at: <http://www.smg.surrey.ac.uk/approaches/canonical-typology/bibliography/>.

make reference to the conceivable extents of dimensions; this means that when observations from the real world are compared to them the results are capable of readily suggesting empirical hypotheses about as-yet unobserved linguistic systems. A characteristic of existing work in canonical typology is the following up of these leads, to more fully reveal the extent of observable typological variation (Aronoff 2019: 138; Nichols 2019).

1.6 Canons, typologies and explanatory theories

The core points of Section 1 can be summarized by contrasting the notion of a canon (Section 1.5) with two other pillars of linguistic methodology: typologies and explanatory theories. A canon is not a typology. Inherently, it is not even a category in a typology (though a multidimensional typology under some circumstances may end up containing a category that matches the canon, just as a unidimensional typology may contain a point that matches a single reference point). Rather, a canon is a methodological device. It is a set of reference points, which aids us in making multidimensional comparisons; this helps us to uncover the underlying attributes of a domain space, and can assist us in selecting dimensions for a typology. A canon can serve as an aid to substruction, by requiring the typologist to identify multiple attributes, and to find reference points such as idealized extremes. A canon can be formulated so as to stand in correspondence to existing knowledge, and thus ensure our typologizing engages with it. A canon is not an explanatory theory. It does not attempt to explain why languages are the way they are, rather it is an aid for a rigorous mode of comparing and organizing empirical data, which can lead to empirical hypotheses and – if it employs idealizations as its reference points – to the identification of gaps between what is conceivable and what is observed. Appreciation of those gaps provides the basis that can lead us from observation to explanation.

1.7 A note on standardization

Standardization is the collective sociological agreement within a community of practitioners to use methods and terminology similarly. By assisting a community to all pull in the same direction, standardization can improve the effectiveness of communication and the transparency of scientific results. In this way, linguists can now profit from several sets of standardized abbreviations,

symbols and data formats.¹¹ However, the benefits of all pulling in the same direction depend upon the direction one is pulling in. Whereas abbreviations and formatting are well suited to effective standardization, the standardization of typological domains and dimensions raises a raft of challenges.¹² Here we focus on one. A deliberate emphasis of this paper is that typically in linguistics we are still attempting to understand the very nature of the phenomena we are investigating, and are routinely discovering new ways in which they vary (Evans and Levinson 2009). In terms of domains and dimensions, every new research project is likely to turn up new directions in which we collectively might decide to pull. In such circumstances, barring extremely good luck, it is reasonable to expect that most standardized definitions and metrics, if they were agreed upon, would soon become obsolete for most purposes. Standardization offers rigidity, which in the right circumstances can be helpful, but at the current stage of linguistic research, for most tasks we need agility. Accordingly, here we have chosen to concentrate on how linguists can best harness that agility to meet the challenge of understanding linguistic diversity.

2 Measurement for comparison, within and across languages

Let us now apply the ideas of Section 1, and also narrow our focus from the general notion of comparison to the specific role of measurement. In Section 2, we illustrate what measurement can look like in linguistics, and examine how it works and what it does for us. Given one recent debate in typology (Haspelmath 2010; Newmeyer 2010), it is notable that our approach to measurement allows one to conduct comparisons equally well within languages and between them, and even to compare within-language variation directly to cross-linguistic variation. This follows from an approach in which we measure the variability, say, of the length of

11 The Leipzig glossing rules provide a list of standardized abbreviations (though not definitions) for widely-used concepts in morphosyntax (Comrie et al. 2008). ISO 639-3 (https://iso639-3.sil.org/code_tables/639/data) and Glottolog (Hammarström et al. 2019) provide distinct codes for languages. Corbett (2013) provides conventions for setting out morphological paradigms. A move towards the harmonization of journal stylesheets can be found at <https://www.eva.mpg.de/lingua/pdf/GenericStyleRules.pdf>. And Forkel et al. (2018) is a proposal for standardized linguistic datasets in digital form.

12 See also Simpson (1999) for problems that can arise for typology when a standardized menu of labels (the International Phonetic Alphabet (International Phonetic Association 1999)), is misconstrued as a cross-linguistic standardization of the diverse, language-specific labelled objects themselves (i.e. phonemes).

vowels, or the range of the genitive case value, using carefully defined criteria (Corbett 2012), a notion which corresponds directly to the ‘dimensions’ mentioned in Section 1. We employ these criteria similarly when comparing the idiolects of two Russian speakers, when comparing older speakers with younger speakers, Russian speakers with Polish speakers, and Polish speakers with Archi or Bezhta speakers. The reason why this is even feasible is that the underlying attributes, which permit variation both within and across languages, are the same, and our criteria (or dimensions) have been constructed carefully so as to reflect those shared attributes. Now, one *could* construct language-internal dimensions in such a way that they fail to generalize cross-linguistically, but it would be a logical error to conclude from this fact, that *all* dimensions which function within one language will necessarily fail to generalize cross-linguistically. Below we illustrate several dimensions that generalize perfectly well. For linguistic typologists, this is heartening news, since it indicates that multi-purpose typological dimensions can indeed be established without insurmountable difficulties, provided we understand that this is what we want to aim for.

The second point we emphasize in Section 2 is the value of setting up measurements so as to operate along a scale. To use an analogy with the measurement of temperature, once we carefully establish scales of measurement, which correspond to the extents of our typological dimensions, then we are no longer confined to labelling our objects of study as just “hot things” or “cold things” (as if the world were that simple); rather we can explore their finer-grained nature, compare it accurately, and report it in detail to our peers. Desirable consequences follow. Variability and empirical uncertainty become easier to characterize, and so can be incorporated into our analyses, rather than factored out of them (see Round 2017 for discussion). And as we measure more carefully, additional analytical tools become available to us, as we participate increasingly fully in mainstream (social) science (Bickel 2015). Thus, in our view linguistics is similar enough to allied fields that – in order to make comparisons – we do not need exceptional devices, but we do need to measure.

2.1 Why measure and how

A great name in measurement is Daniel Gabriel Fahrenheit (1686–1736); he tends to be remembered for a quaint temperature scale, yet his real achievement was truly significant. He “stunned the world by making a pair of thermometers that both gave the same reading” (Lienhard 2000: 162, and see Grigull 1986 for an assessment of Fahrenheit’s stature). Before Fahrenheit, two researchers could not know they had comparable readings in terms of temperature. So Fahrenheit should be

remembered for making measurement possible in a particular domain. To compare we need to be able to measure different things, even when they are not in the same place, the same laboratory or indeed the same language. And crucially we need to get the same results.¹³

At the same time, we should note issues with Fahrenheit's scale, which also prove instructive here. Fahrenheit based his scale on the temperature of freezing brine (his zero), freezing water and the human body (not quite accurately at 96 degrees), but since objects can and do reach temperatures even colder than freezing brine, the effect is to place 'zero' at an artificial, non-extreme centre. The scale then extends outwards on both sides of zero, with positive and negative values. In typology we sometimes make the same mistake, giving undue significance to particular, salient points, by placing them at the centre of our scale and measuring on both sides. Instead, for reasons discussed in Section 1.5, we would do best to anchor our scales at idealized extremes, as Kelvin notably did with his temperature scale anchored at the lowest possible thermodynamic temperature, absolute zero.

2.2 Measuring the Russian genitive

What could we measure in the few pages allowed here? We will use measurement to answer (in small part only) two searching questions due to Andrey Kolmogorov:

What exactly do we mean when we say that two words are in the same case?
How many cases does the Russian language possess?

These are great questions, put to linguists in 1956, by a mathematician (he is famous for Kolmogorov complexity).¹⁴ Let us take one of the easier parts,

13 Some may think we should not measure, but should stick with our intuitions. Other disciplines have faced similar challenges, and the result has been that: "Gradually, however, each such objection has fallen to the inexorable march of measurement. The defeats have not come about through careful arguments convincing the doubters, but simply because of what could be achieved through measurement." (Hand 2016: 115)

14 They were keynote questions put to a seminar on mathematical methods in linguistics, which began in September 1956 at the Philological Faculty of Moscow University; Kolmogorov was not himself present. See van Helden (1993: 138) for an account and an analysis of the significance of the Set-theoretical School, which resulted from the seminar. Various researchers have attempted to answer Kolmogorov's questions; see Corbett (2012: 200–222) for sources and a new attempt; the answer to the second question lies between 6 and 10 and is probably not a whole number.

concerning the case values of Russian. It is regularly stated that Russian has a genitive, one which fits typical definitions.¹⁵ Examples of the Russian genitive are easy to find:

Russian

- (1) *knig-a* *otc-a*
 book-SG.NOM father-SG.GEN
 ‘father’s book’

What is there to measure here? Our simple example already shows that given an *-a* inflection, we do not know that we have a genitive. Although *-a* is indeed a genitive marker, it is not uniquely genitive: see the first noun *kniga* ‘book’, where the *-a* signals something else (a nominative). Moreover, the genitive is not always realized as *-a*:

- (2) *knig-a* *mater-i*
 book-SG.NOM mother-SG.GEN
 ‘mother’s book’

Even given a noun that does realize the genitive with *-a*, this *-a* may not be uniquely genitive:

- (3) *Ja* *viž-u* *otc-a*
 1SG see-1SG father-SG.ACC
 ‘I see father’

From these observations, we might hypothesize that in Russian the genitive and accusative are not distinct. But this is quickly dispelled, since for other nouns the genitive and accusative are distinct (compare (4) with (2) above):

- (4) *Ja* *viž-u* *mat’*
 1SG see-1SG mother[SG.ACC]
 ‘I see mother’

¹⁵ First, a general definition: “Case whose basic role is to mark nouns or noun phrases which are dependents of another noun” (Matthews 1997: 144). And in a grammar of Russian: “Possessors that are nouns are expressed in the genitive, and are placed after the possessed noun ...” (Timberlake 2004: 205). Timberlake expands this with: “As has long been observed, possession should be understood very broadly, ...” (2004: 206). For recent discussion of the definition of genitive see Lander (2009); for other adnominal case values, notably the proprietive, see Dench and Evans (1988); and for helpful discussion of the terminology of case more widely see Haspelmath (2009).

Already it appears that we have a genitive, as the tradition has it. However, when we recall Kolmogorov's first question and apply it to the genitive (asking what it means to say that different words are in the genitive), we find there are in fact many factors in play, all at once. This situation arises often in linguistics and other typological disciplines and so it is valuable to have an approach that allows us to proceed methodically.

2.3 A dimension of comparison, and an extreme point of reference

There is indeed something to be investigated here, and most likely something multidimensional, so what dimension should we measure first? And what would be our zero, the extreme end point of the dimension, from which we measure? In Section 1.1 we mentioned that the attributes that underlie typological dimensions are sometimes shared by multiple domains, noting that in instances where they are, it may be possible to synthesize typological results across those domains. Conversely though, shared attributes can also give us a head start when looking at new domains. So, is there a canonical domain, previously investigated, from which we might borrow a dimension to assist us in measuring the genitive? There is. A useful canon is the canon for inflectional feature values (Corbett 2015: 149–158; Stump 2017). In the extreme, idealized, canonical world, a case value (indeed any feature value) shows a unique mapping, from form to (grammatical) meaning and from meaning to form. Specifically:

- one form one meaning: *-a* (or whatever) realizes genitive only
- one meaning one form: genitive is realized by *-a* only

We are not saying that the real world should match the canonical idealization, any more than that objects “should” have zero length and be at 0 K. We are establishing points to measure from. As Stump nicely describes it, an inflectional paradigm that is canonical ‘is like the winter solstice – a well-defined seasonal extreme relative to which a year’s 364 other days can be calendrically compared, but by no means the most common or the most typical of days’ (2016: 103). What we are saying is that we establish a canon to measure from, which is preferable to measuring arbitrarily and ambiguously from the situation in Standard Average

European, or from a prototype established in an individual language or set of languages.¹⁶

How does the Russian genitive measure up against this canon? Let us look at the fuller picture in Table 1.

Table 1: Russian nouns illustrating the inflectional status of the genitive.

	'father'		'mother'		'magazine'	
	SG	PL	SG	PL	SG	PL
NOMINATIVE	otec	otcy	mat'	materi	žurnal	žurnaly
ACCUSATIVE	otca	otcov	mat'	materej	žurnal	žurnaly
GENITIVE	otca	otcov	materi	materej	žurnala	žurnalov
DATIVE	otcu	otcam	materi	materjam	žurnalu	žurnalam
INSTRUMENTAL	otcom	otcami	mater'ju	materjami	žurnalom	žurnalami
LOCATIVE	otce	otcax	materi	materjax	žurnale	žurnalax

Table 1 illustrates the inflectional behaviour of three different classes of Russian nouns. We see that neither *otec* 'father' nor *mat'* 'mother' has a unique genitive form. However, there are nouns like *žurnal* 'magazine', which have forms unique to the genitive both in the singular and the plural. Consider now

16 Mel'čuk has pressed for conceptual and terminological clarity. We have learnt from his work, but we part company over his "orientation towards prototypical cases" (2006: 175–177). This move is intended to maintain sensible use of terms. However, while prototypes may be appropriate for speakers, they are less good for analysts; they encourage us to see less prototypical instances in terms of the prototype (as we suggested in Section 1.5).

As an example, consider inflectional classes. We might consider Latin, Spanish or Russian as providing the prototype here. Indeed, discussion in the literature has focussed on such prototypical systems, with their familiar clustering of properties. Contrast that with a canonical definition which takes the essential characteristics of inflectional classes to their logical end point. There are two principles: first 'distinctiveness': 'Canonical inflectional classes are fully comparable and are distinguished as clearly as is possible'; and second 'independence': 'The distribution of lexical items over canonical inflectional classes is synchronically unmotivated'. The principles are spelled out in a set of more specific criteria. One might reasonably assume, as we did, that no clear case of canonical inflectional classes would exist; rather they would remain an unrealized idealization. The next stage would be to look for reasons why such systems cannot exist.

But it turns out that they can. Burmeso (bzu, Donohue 2001, discussed in Corbett 2009) is an almost perfect example of canonical inflectional classes. For instance, Burmeso verbs have two inflectional classes and absolutely no forms are shared between them. Thus our approach leads to better results: we do not see Burmeso through the lens of prototypical Latin inflectional classes. We start from a more abstract measure. We recognize the special interest of Burmeso. And we can ask why such systems are rarer than that of Latin.

what this means for the question of whether or not any two given nouns in Russian are in the same case (and in particular, whether or not they are both in the genitive). We might propose, looking at just the forms in Table 1, that only certain Russian nouns, like *žurnal*, truly have a genitive at all, while *otec* ‘father’ and *mat’* ‘mother’ simply lack one. The claim would be, that in the syntactic situations where *žurnal* appears in the genitive, *otec* and *mat’* appear in some other case. However, the serious problems caused by this kind of analysis are well established and we will not rehearse them here (see Zaliznjak 1973 or Blake 1994: 20–23 for the arguments). Rather, following the definition of our domain, we wish to analyse syntactically equivalent ‘dependents of another noun’ as standing in the equivalent case to one another (the genitive); this holds even when their morphology is ambiguous. All nouns have a genitive, irrespective of morphological differences such as those between *otec*, *mat’* and *žurnal*. Even so, while the statement that all Russian nouns have a genitive is true, this is not the whole picture. Many nouns do not have a unique genitive form. And whether they do or do not is something we can measure, noun by noun. Given Zaliznjak’s (1977) grammatical dictionary of Russian, we find that of the 37,678 nouns included, as many as 24% lack a unique genitive form, even when stress alternations as well as segmental phonology are taken into account (thanks to Dunstan Brown for this calculation). This is perhaps not what the reader expects, having learnt previously from grammatical descriptions that Russian ‘has a genitive’ among its case values; this new characterization, thanks to measurement, is more precise.

We now present these initial data graphically (Figure 1). In this instance, it makes sense to express our measure numerically as a proportion, with values between zero and one. As a convention, we place the canonical value at zero, which in this instance means that maximally non-canonical corresponds to 1. (This

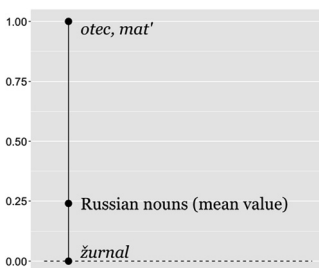


Figure 1: Formally unique genitive: measuring within Russian. Zero is canonical. Here, maximally non-canonical is 1.

will not always be the case. Other measures could equally well extend outwards from zero with no upper limit, while yet others could be categorical, for example splitting observations simply into ‘canonical’ versus ‘noncanonical’: see Figures 4 and 6 below.)

In Figure 1, the result would be the canonical zero if every Russian noun were like *žurnal* ‘magazine’, in having a unique genitive; it would be non-canonical ‘1’, if all were like *otec* ‘father’, and *mat’* ‘mother’ in lacking a unique genitive. Counting across the noun lexicon shows that some 24% do not have a unique genitive, that is, we are .24 of the way from canonical towards non-canonical.¹⁷

2.4 Comparison by measuring relative to the canon, both cross-linguistically and internally

Now let us contrast Russian with the noun paradigm of the Dagestanian language Archi (aqc) in Table 2; we omit the many spatial case values (see Chumakina et al. 2007: vi, and sources there).

Table 2: Archi cases (non-spatial).

‘ram’	SINGULAR	PLURAL
ABSOLUTIVE	ba ^ʕ kʰ	ba ^ʕ kʰ-ur
ERGATIVE	be ^ʕ kʰ-iri	ba ^ʕ kʰ-ur-čaj
GENITIVE	be ^ʕ kʰ-iri-n	ba ^ʕ kʰ-ur-če-n
DATIVE	be ^ʕ kʰ-iri-s	ba ^ʕ kʰ-ur-če-s
COMITATIVE	be ^ʕ kʰ-iri-ɬ:u	ba ^ʕ kʰ-ur-če-ɬ:u
SIMILATIVE	be ^ʕ kʰ-iri-q ^ʕ di	ba ^ʕ kʰ-ur-če-q ^ʕ di
CAUSAL	be ^ʕ kʰ-iri-š:i	ba ^ʕ kʰ-ur-če-š:i
COMPARATIVE	be ^ʕ kʰ-iri-χur	ba ^ʕ kʰ-ur-če-χur
PARTITIVE	be ^ʕ kʰ-iri-q ^ʕ iš	ba ^ʕ kʰ-ur-če-q ^ʕ iš
SUBSTITUTIVE	be ^ʕ kʰ-iri-kʰʼena	ba ^ʕ kʰ-ur-če-kʰʼena

¹⁷ We have restricted the measure here to canonical (unique) versus non-canonical (not unique), and then measured across the noun lexicon. But the method offers a finer-grained measure. As Table 1 indicates, there is a further distinction that can be drawn between *otec* ‘father’ and *mat’* ‘mother’: the genitive of the former is syncretic with one other case value only (the accusative), while the genitive of *mat’* ‘mother’ shows further syncretism, and so is less distinct. For simplicity, we have stayed with the basic measure.

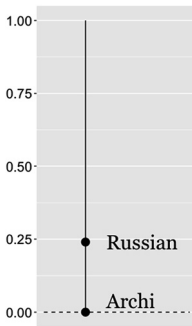


Figure 2: Formally unique genitive: measuring Russian and Archi nouns.

First, *-n* is the marker of the genitive for every noun in Archi. This allows a clear contrast with Russian (See Figure 2).

While we have concentrated on nouns, we should mention pronouns too. The Archi pronouns also have unique forms for the genitive, but they are not identical (that is, each pronoun has a distinct genitive form, but the marker is not segmentally the same for all pronouns). We should compare the Russian pronouns (See Table 3).

We see, perhaps surprisingly, that in Russian none of the personal pronouns has a unique genitive form. Nevertheless, the distributional argument given above means that we still recognize a genitive here. Returning to comparison, note that

Table 3: Russian personal pronouns.

	1		2		3			
	SG	PL	SG	PL	SG M	SG F	SG N	PL
NOMINATIVE	ja	my	ty	vy	on	ona	ono	oni
ACCUSATIVE	menja	nas	tebja	vas	ego	ee	ego	ix
GENITIVE	menja	nas	tebja	vas	ego	ee	ego	ix
DATIVE	mne	nam	tebe	vam	emu	ej	emu	im
INSTRUMENTAL	mnoj	nami	toboj	vami	im	ej/eju	im	imi
LOCATIVE	mne	nas	tebe	vas	nem ^a	nej	nem	nix

^aWhen directly governed by a preposition, third person pronouns take a prothetic *n-*. The locative is found only when governed by a preposition, hence we include the *n-* for the locative.

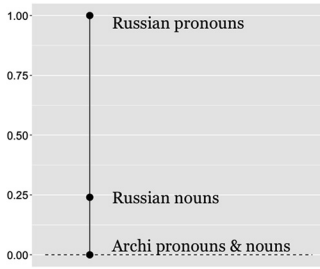


Figure 3: Formally unique genitive: measuring Russian and Archi nouns and pronouns.

we compare nouns and pronouns within Russian exactly as we compare between Russian and Archi (See Figure 3).¹⁸

2.5 Additional dimensions and their canonical extremes

Let us now bring in a second typological dimension along which we can measure and compare the genitive. If we return to the Archi paradigm in Table 2, we see that Archi marks the genitive uniformly across number: it is the same in the singular and the plural. In Russian this is not the case. Thus, when viewed across the values of orthogonal features (i.e. singular and plural number), the form-meaning mapping is clearly closer to unique in Archi than in Russian. In Figure 4 we show this in

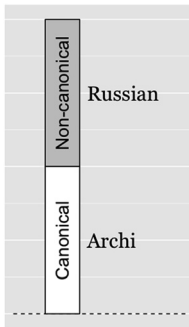


Figure 4: Formally unique genitive across features: measuring Russian and Archi nouns and pronouns.

¹⁸ For recent instances of using canons for language-internal analysis see Kwon (2017) and Menzel (2018).

terms of a simple, categorical measure. It would also be possible to operationalize the notion of ‘unique across features’ in more fine-grained terms, but for reasons of space, we will not do that here.

There is more to be measured. Let us switch from ‘father’s book’ (as in (1) above) to ‘Dad’s book’:

Russian

- | | | | |
|-----|--|---------------------------------|------------------------|
| (5) | <i>knig-a</i>
book-SG.NOM
‘Dad’s book’ | <i>pap-y</i>
Dad-SG.GEN | (genitive) |
| (6) | <i>pap-in-a</i>
Dad-POSS-F.SG.NOM
‘Dad’s book’ | <i>knig-a</i>
book(F)-SG.NOM | (possessive adjective) |

We see that in Russian, alongside the normal genitive in (5), there is an alternative: in (6) we have a possessive adjective (which agrees in gender and number with the head noun). This is another way in which the genitive of a language may be unique or not. The noun *papa* ‘Dad’ has a competitor for the genitive slot, in the form of the derived possessive adjective (Corbett 1995).¹⁹ In other languages we are used to seeing such competition, but more usually with pronouns. Thus for the dependent of another noun, we are used to English *my* rather than *of me*, and similarly in Russian we have: *moja kniga* ‘my book’, not **kniga menja* ‘book of me’. But in (6) we see that competition extends to nouns too (though it is limited as to where it can occur; for instance, it is not readily available with all nouns, and it is restricted to the singular of the possessor).

This competition reveals a distinct way in which the Russian genitive is not unique. Thinking typologically, it suggests another distinct dimension along which we can attempt to observe variation and a corresponding, distinct type of measurement. This is fine; physicists do not use only temperature after all. In order to measure this second kind of uniqueness along a scale, we can count the examples of possessive adjectives in use, as opposed to the genitive. Fortunately there are statistics both for Russian, and for some of the related Slavonic languages, allowing us to relate to a limited cross-linguistic typology (Section 1.4). These statistics are not as detailed as we might hope, but they show something of the range of variation. Ivanova (1976: 9–10) gives figures based on contemporary literature, criticism and journalism. For each language investigated she examined

¹⁹ A rich and extensive survey of the strategies of adnominal possession in the languages of Europe is provided by Koptjevskaja-Tamm (2003), Nikolaeva and Spencer (2013), and Spencer (2013: 348–356) offer a recent insightful discussion of possession and modification, laying out the possibilities in a typologically diverse range of languages.

1,000 pages (counting 2,000 characters as a page). She gives the approximate frequencies of use of the possessive adjective,²⁰ expressed as a percentage of the total instances of the possessive adjective and of the genitive (without preposition) in the first column of Table 4. These figures furnish us with measurements that illustrate the difference in the balance of the competition between genitive and possessive adjective in related languages (bear in mind that in some instances the possessive adjective would be excluded by the restrictions discussed above).

Table 4: Frequency of use of the possessive adjective.

	Overall	In situations of choice only
East Slavonic		
Russian (rus)	10%	22%
Belarusian (bel)	36%	65%
Ukrainian (ukr)	23%	49%
South Slavonic		
Slovene (slv)	66%	98%
Serbo-Croat (hbs)	52%	93%
West Slavonic		
Polish (pol)	3%	6%
Czech (ces)	51%	94%
Slovak (slk)	42%	83%

Looking first at the Russian figure of 10% we see that overall the normal genitive (as in (5) above) is much more frequent than the possessive adjective (6). However, there is considerable variation within the family, with the South Slavonic languages favouring the possessive adjective. Ivanova (1975: 151) provides more useful figures, in that she counted just those instances where the use of the possessive adjective is theoretically possible (for a singular referent, with no modifier in the corresponding genitive phrase,²¹ and not expressed by an adjectival noun – which could not form a possessive adjective). These data are given in column 2 of Table 4. Naturally, the possessive adjective achieves a higher frequency under these conditions, but the differences between the languages investigated are equally clear. They are summarized in Figure 5, where the canonical situation for the genitive is that it is the unique possibility.

²⁰ The number of actual instances is not given by Ivanova; we have rounded her figures to whole percentages.

²¹ This restriction does not apply fully to Slovak, but Ivanova does not specify how she handled this.



Figure 5: Unique genitive (compared with competing form).

There is a further way in which the genitive may not be unique. In Russian, the genitive is used for the dependent of another noun (as in Matthews' definition) irrespective of the case of the head. But this is not the only possibility, as these data from the Dagestanian language Bezhta (kap) demonstrate (Kibrik 1995: 220, van den Berg 2005: 261; further examples in Boguslavskaja 1995: 233–234):

Bezhta

- (7) *abo-s* *is*
 father-GEN_{DIRECT} brother[ABS]
 'father's brother'
- (8) *abo-la* *is-t'i-l*
 father-GEN_{OBLIQUE} brother-OBL-DAT
 'to father's brother'

Example (7) has a genitive as expected. However, this genitive is available only if the head is in the absolutive case. In all other instances, a different case form (here labelled genitive oblique) is used, as in (8). All our previous examples were more canonical than those of Bezhta, in that they involved a genitive, that is unique in the sense that it was not differentially conditioned by the case of the head (see Figure 6).



Figure 6: Unique genitive (syntactically).

2.6 What canonical analysis can reveal

Having established several typological dimensions for characterizing variation within our domain (though by no means having exhausted all possibilities), we can now present the overall picture gained from our multiple measurements in Figure 7, and make a few, selected remarks.

Seeing the measurements side by side, partial though they are, already prompts interesting typological conclusions.²² For instance, differences within a language can be greater than those across languages (thus the first scale in Figure 7 shows that Russian nouns in general are closer in behaviour, in one respect, to Archi nouns and pronouns than they are to Russian pronouns). And more generally, a language may show canonical behaviour in one respect and highly non-canonical behaviour in another.²³ This way of characterizing the

²² And recall that we have tackled just one of the easier parts of Kolmogorov's questions. Moreover, in addition to Kolmogorov's concerns, and so properly outside our topic here, there is much more that could be said about the Russian genitive. First there is its competition with the receding second genitive (Corbett 2012: 203–206). There are interesting questions about the available range of meanings for the adnominal genitive (extensive in Russian). Then there is its prominent role in quantified expressions, a role which has called forth a veritable tidal wave of ink. And then, the significant fact that it is not restricted to adnominal use, but it also operates at clause level: it is used for the object of certain transitive verbs (such as *bojat'sja* 'fear'), and most notably as a case for negated transitive verbs (in competition with the accusative). There are numerous conditions on the accusative/genitive choice; for sources, and an online searchable database see Krasovitsky et al. (2009). These additional dimensions for understanding the genitive are all amenable to a canonical approach; the issue of the accusative/genitive choice in particular has been subject to careful measurement in corpora, for example in Mustajoki and Heino (1991).

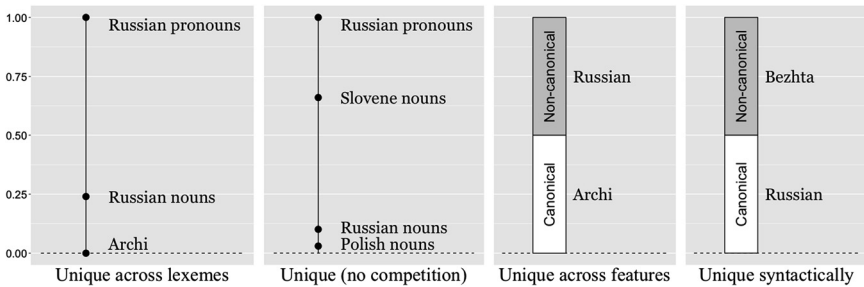


Figure 7: The Russian genitive in typological context.

data, in a multidimensional fashion, is more informative and more specific than reducing all variation down to a single value (Cartwright and Bradburn 2011: 63), for example by collapsing rich and complex distinctions into a simple, monolithic cline, such as by measuring a one-dimensional ‘distance’ of a language from a prototype.²⁴

3 Conclusion: Linguistic typology, a typological science in good company

The science of language is uniquely interesting and highly challenging. Yet there are pioneers in other disciplines who have tackled phenomena of similar complexity for which the categories, units, and scales of measurement are also far from settled and uncontentious. We should take advantage of their experiences and methods, since the daunting challenges they face are essentially the same as ours. Progress is made by

23 An analogy may prove helpful here. Suppose we measure a set of objects in terms of temperature (from 0 K) and length (from 0 m). There is nothing in the two scales which implies that there should be any correlation between the two; measuring in this way does not imply that larger objects will be hotter; and smaller objects colder. We *might* have reasons to suggest a connection, but that is a matter of observation, analysis and theorizing.

24 More sophisticated analyses that can be profitably applied to multidimensional data include factor analysis (advocated for typological sciences already by Winch 1947), principal component analysis, and related methods, though ideally with careful attention to matters of phylogenetic non-independence (Bromham, this issue).

examining complex phenomena along multiple typological dimensions, a viewpoint shared by Canonical Typology and Multivariate Typology (see further in Appendix). In Canonical Typology, we measure along these multiple dimensions, and where possible define scales of measurement relative to idealized extremes, such as zero. We do this in linguistics for the same reason as in other disciplines: because points of reference are fundamental tools for establishing measurements, and because idealized extremes avoid inherent problems associated with alternative points of reference (such as prototypes and standards). The results that we obtain with their help are of interest whether or not any real-world instances actually meet these idealized extremes. Idealized extremes, including multi-dimensional canons, are not data or explanatory theories; they are a non-arbitrary, low-ambiguity means of measuring what is observed and comparing it to what is conceivable, an essential step towards explaining what is possible.

In any typological science, progress is hard-won. Often, it is less rapid than we desire. Such is life, but we have grounds for optimism in linguistic typology. For over a century, often in the form of independent, parallel discoveries, leading thinkers in typological fields have emphasized the same core set of methodological techniques whose utility we have emphasized here. All typological proposals require debate that leads to refinement, and the productivity of debate can be enhanced with substruction. By relating our typological dimensions to the underlying attributes of a domain (whatever we take those attributes to be, and however murky they currently are), we can explain to ourselves and our peers the motivation for our choices. The same goes for operationalization and the parcelling out of dimensions. By doing this we sow the seeds of more lucid and productive debate, thus more thorough evaluation of competing proposals, which means progress.

Comparison can be carried out similarly within and across languages. To compare across Russian, Polish, Archi, and Bezhta we can use exactly the same measures as for comparing within Russian or between different individual speakers. This has significant implications for our discipline. Namely, we should be careful of erecting a methodological framework based on the notion that somehow language-internal and language-external comparisons are inherently different (pace Haspelmath 2010).²⁵ Linguistics faces the same core challenges that

²⁵ Haspelmath's diagnosis of what linguists can hope to know about language is the subject of an extended analytic examination by Spike (this issue), who reveals it to be rooted in a rather singular philosophical understanding of science. We find Spike's analysis enlightening, since it pinpoints where we and Haspelmath ultimately diverge in our views; and in some senses a relief, since it

characterize any typological field. By eschewing radical responses that would paint linguistics as somehow unique, we can ensure that comparison in linguistics continues to be conducted as in other typological sciences, of which there are many. Doing so enables us to remain within the scientific mainstream (Bickel 2015), where we can reasonably expect to continue to profit from and contribute to allied fields. This is much more promising than isolation.

To conclude: Typologists have been obtaining exciting results for decades, by applying similar measures to language-internal and cross-linguistic comparison. This is the right course. Here we have provided reasons to stay on it, and have detailed how best to do so.

reassures us that we are most likely on the right side of the argument. Simplifying greatly (and we urge readers to see Spike for the full analysis), Haspelmath would appear to assume that what we have treated as real-world attributes of linguistic phenomena – attributes which cross-linguistically are exploited by languages in many varied ways, and which can be employed by typologists to build dimensions of comparison within and across languages – are, for philosophical reasons, simply not accessible to linguistic science and perhaps do not even exist. Consequently, to compare across languages linguists must create an alternative to the ineluctably intractable real world by inventing their own, purely instrumental categories, termed ‘comparative concepts’. If this is true, then it would be something of a grim world in which to conduct linguistic typology. Substruction could not exist, since without an accessible, external real world to appeal to, linguists have nothing concrete against which to assess their proposed dimensions. Debate over categories perhaps could persist, at best as an intellectual exercise and at worst as a power struggle for arbitrary conceptual hegemony, but being untethered to an external reality, there is nothing ‘out there’ that typologists could collectively appeal to or even aspire to discovering. Thus, Haspelmath’s programme seeks to define concepts by fiat, thereby ending debate over them, and to wall off language-particular analysis from cross-linguistic research. And yet we find it conspicuous that other typological disciplines with longer and richer histories of philosophical introspection have not arrived at the same conclusion (see Spike on why this is so). Our own position, and we suspect the position of most linguistic typologists, is consistent with a more mainstream understanding of the ways and possibilities of science in general. We regard linguistics to be among the typological disciplines. Its task is discovery and refinement through productive debate. Language-internal and cross-linguistic research can be mutually informative. Accordingly, there is no need for radically ungrounded comparative concepts *sensu* Haspelmath, nor for the separation of linguistic typology from linguistic description. Such ‘cures’ for typology are unnecessary. They are responses to an illusory malady, that only appears to exist under one peculiar and, we believe, ultimately erroneous epistemological outlook. Returning to concrete matters, since these cures demand a near complete re-building of our discipline; and the hard isolation of it from much prior research and related, fascinating enterprises within and beyond linguistics, they come with a weighty opportunity cost. Quite simply, in our view, there are more profitable goals for the discipline to pursue; we are in good health, so let’s get on with research. The discovery-based, substructive methodology outlined here and employed equally by other sciences facing similar challenges to our own, is the most promising, practical, and optimistic choice for the field.

Acknowledgement: We are thankful to the Centre of Excellence for the Dynamics of Language for bringing Corbett to the ALT special session in Canberra and facilitating our ongoing collaboration. The support of the AHRC (UK), grant AH/N006887/1: Lexical splits: a novel perspective on the structure of words, the British Academy Visiting Fellowships Programme under the UK Government's Rutherford Fund, grant VF1_101602, and the ARC (Australia), DECRA award DE150101024, is gratefully acknowledged. We thank Nick Evans, our co-contributors Lindell Bromham and Matthew Spike, together with Sebastian Fedden, Tim Feist, Steven Kaye and Terje Lohndal, for detailed constructive comments on various drafts, and Balthasar Bickel, Olivier Bonami, Martin Haspelmath and Andrew Spencer for helpful discussion at different times. All responsibility for the views and arguments herein remains our own.

Appendix: Canonical typology and multivariate typology

While a thorough comparison would take us beyond the scope of this paper, we have been asked to compare briefly Canonical Typology (CT) with Multivariate Typology (MT), as set out, for instance, in Bickel (2007, 2015). These two approaches to typology both arose in the early 21st century. They share an essential outlook, which is that linguistic variation is fundamentally multi-dimensional, and possesses a level of nontrivial detail that far surpasses what previous generations of typology sought to examine. Both approaches regard a core task of linguistic typology to be the careful characterization and discovery of these dimensions of variation, and view them as processes that operate hand in hand with wide-ranging comparison of real-world phenomena. As we see it, there is also a cluster of mutually related differences. In practice, research in MT has focussed on *attested* data points, and for the most part has not set out to establish the conceivable extents of dimensions. In CT, idealized reference points and conceivable extents have a privileged place in the methodology. A corollary is that empirical research in CT tends to employ directed searches for specific linguistic properties, as one attempt to confirm empirical hypotheses about as-yet unobserved phenomena that have been postulated on the basis of the conceivable. As a result, CT tends to contribute particularly to the expansion of the outer bounds of confirmed linguistic diversity. Relatedly, CT is associated with the creation of databases whose languages are sampled with the aim of showcasing the *range of possible variation* within a domain. MT on the other hand has been employed particularly in building numerically larger databases that reflect the *distribution of*

attested variation within some sample of languages, chosen according to prior criteria. Correspondingly, MT contributes particularly to our understanding of the relative frequencies of typological phenomena and their correlations with genealogy and extra-linguistic matters such as geography and demography. The fact that the two research programs have been distinct appears to us more an accident of history and a matter of different ranking of research priorities, rather than any underlying incompatibility or contradictions in the aspirations of the ultimate research program. See also Forker (2016) for a fuller comparison of CT and MT.

References

- Aronoff, Mark. 2019. Canonical syncretism and Chomsky's S. In Matthew Baerman, Oliver Bond & Andrew Hippisley (eds.), *Morphological perspectives*, 138–147. Edinburgh: Edinburgh University Press.
- Bickel, Balthasar. 2007. Typology in the 21st century: Major current developments. *Linguistic Typology* 11. 239–251.
- 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 edn., 901–923. Oxford: Oxford University Press.
- Blake, Barry J. 1994. *Case*. Cambridge: Cambridge University Press.
- Boguslavskaja, Ol'ga Ju. 1995. Genitives and adjectives as attributes in Daghestanian. In Frans Plank (ed.), *Double case: Agreement by Suffixaufnahme*, 230–239. New York: Oxford University Press.
- Bond, Oliver. 2013. A base for canonical negation. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 20–47. Oxford: Oxford University Press.
- Bond, Oliver. 2019. Canonical typology. In Jenny Audring & Francesca Masini (eds.), *The Oxford handbook of morphological theory*, 409–431. Oxford: Oxford University Press.
- Cartwright, Nancy & Norm Bradburn. 2011. A theory of measurement. In Rose Maria Li (ed.), *The importance of common metrics for advancing social science theory and research*, 53–70. Washington DC: National Academies Press.
- Chumakina, Marina, Dunstan Brown, Harley Quilliam & Greville G. Corbett. 2007. *Slovar' arčinskogo jazyka (arčinsko-russko-anglijskij) [A dictionary of Archi: Archi-Russian-English]*. Makhachkala: Delovoj Mir.
- Comrie, Bernard, Martin Haspelmath & Balthasar Bickel. 2008. *The Leipzig glossing rules: Conventions for interlinear morpheme-by-morpheme glosses*. Leipzig: Department of Linguistics of the Max Planck Institute for Evolutionary Anthropology & the Department of Linguistics of the University of Leipzig. <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>.
- Corbett, Greville G. 1995. Slavonic's closest approach to Suffix Copying: The possessive adjective. In Frans Plank (ed.), *Double case: Agreement by Suffixaufnahme*, 265–282. New York: Oxford University Press.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2009. Canonical inflectional classes. In Fabio Montermini, Gilles Boyé & Jesse Tseng (eds.), *Selected proceedings of the 6th Décembrettes: Morphology in Bordeaux*, 1–11. Somerville, MA: Cascadilla Proceedings Project. <http://www.lingref.com/cpp/decemb/6/abstract2231.html>.
- Corbett, Greville G. 2012. *Features*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2013. Paradigm conventions. *Paper presented at the 46th annual meeting of the Societas Linguistica Europaea, Split, 18–21 September 2013*. https://www.academia.edu/9055930/Paradigm_conventions.
- Corbett, Greville G. 2015. Morphosyntactic complexity: A typology of lexical splits. *Language* 91. 145–193.
- Corbett, Greville G. & Sebastian Fedden. 2016. Canonical gender. *Journal of Linguistics* 52. 495–531.
- Cormier, Kearsy, Adam Schembri & Bencie Woll. 2013. Pronouns and pointing in sign languages. *Lingua* 137. 230–247.
- Daniels, Don & Greville G. Corbett. 2019. Repartitioning. *Language* 95. 711–750.
- De Queiroz, Kevin. 2005. Different species problems and their resolution. *BioEssays* 27. 1263–1269.
- Dench, Alan C. & Nicholas Evans. 1988. Multiple case-marking in Australian languages. *Australian Journal of Linguistics* 8. 1–47.
- Dennett, Daniel C. 1996. *Darwin's dangerous idea: Evolution and the meanings of life*. London: Penguin.
- Donohue, Mark. 2001. Animacy, class and gender in Burmeso. In Andrew Pawley, Malcolm Ross & Darrell Tryon (eds.), *The boy from Bundaberg: Studies in Melanesian linguistics in honour of Tom Dutton (Pacific Linguistics 514)*, 97–115. Canberra: Pacific Linguistics.
- Ember, Carol R., Marc Howard Ross, Michael L. Burton & Candice Bradley. 1991. Problems of measurement in cross-cultural research using secondary data. *Behavior Science Research* 25. 187–216.
- Evans, Nicholas. 2013. Some problems in the typology of quotation: A canonical approach. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 66–98. Oxford: Oxford University Press.
- Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018a. The grammar of engagement I: Framework and initial exemplification. *Language and Cognition* 10. 111–140. <https://doi.org/10.1017/langcog.2017.21> [open access].
- Evans, Nicholas, Henrik Bergqvist & Lila San Roque. 2018b. The grammar of engagement II: Typology and diachrony. *Language and Cognition* 10. 141–170. <https://doi.org/10.1017/langcog.2017.22> [open access].
- Evans, Nicholas & Alan Dench. 2006. Introduction: Catching language. In Felix K. Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 1–39. Berlin: Mouton de Gruyter.
- Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32. 429–448.
- Fedden, Sebastian & Greville G. Corbett. 2017. Gender and classifiers in concurrent systems: Refining the typology of nominal classification. *Glossa: A Journal of General Linguistics* 2(1), 34. 1–47.

- Feest, Uljana. 2005. Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences* 41. 131–149.
- Forkel, Robert, Johann-Mattis List, Simon J. Greenhill, Christoph Rzymiski, Sebastian Bank, Michael Cysouw, Harald Hammarström, Martin Haspelmath, Gereon A. Kaiping & Russell D. Gray. 2018. Cross-linguistic data formats, advancing data sharing and re-use in comparative linguistics. *Scientific Data* 5. 180205.
- Forker, Diana. 2016. Conceptualization in current approaches of language typology. *Acta Linguistica Hafniensia* 48. 70–84.
- Gerlach, Martin, Beatrice Farb, William Revelle & Luís A. Nunes Amaral. 2018. A robust data-driven approach identifies four personality types across four large data sets. *Nature Human Behaviour* 2. 735–742.
- Greenberg, Joseph. 1974. *Language typology: A historical and analytic overview*. The Hague: Mouton.
- Grigull, Ulrich. 1986. Fahrenheit a pioneer of exact thermometry. *Heat Transfer 1986: Proceedings of the eighth international heat transfer conference, San Francisco, USA, I*, 9–18. Washington: Hemisphere. <http://ihtcdigitallibrary.com/download/article/1480b040272d93c1/pl-2.pdf>.
- Hammarström, Harald, Robert Forkel & Martin Haspelmath. 2019. Glottolog 4.0. Jena. <https://glottolog.org/>.
- Hand, David J. 2016. *Measurement: A very short introduction*. Oxford: Oxford University Press.
- Haspelmath, Martin. 2009. Terminology of case. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 506–517. Oxford: Oxford University Press.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86. 663–687.
- Haspelmath, Martin. 2012. How to compare major word-classes across the world's languages. In Thomas Graf, Denis Paperno, Anna Szabolcsi & Jos Tellings (eds.), *Theories of everything: In honor of Edward Keenan (UCLA Working Papers in Linguistics 17)*, 109–130. Los Angeles: UCLA.
- Hempel, Carl G. & Paul Oppenheim. 1936. *Der Typusbegriff im Lichte der neuen Logik*. Leiden: Sijthoff.
- Hull, David L. 1968. The operational imperative: Sense and nonsense in operationism. *Systematic Biology* 17. 438–457.
- Hyman, Larry M. 2012. In defense of prosodic typology: A response to Beckman & Venditti. *Linguistic Typology* 16(3). 341–385.
- International Phonetic Association. 1999. *Handbook of the international phonetic association: A guide to the use of the international phonetic alphabet*. Cambridge: Cambridge University Press.
- Ivanova, T. A. 1975. Nekotorye aspekty sopostavitel'nogo analiza possessivnyx konstrukcij (Na materiale sovremennyx slavjanskix literaturnyx jazykov) [Aspects of the comparative analysis of possessive constructions (based on contemporary Slavonic literary languages)]. *Slavjanskaja filologija* 3. 148–152. (Leningrad).
- Ivanova, T. A. 1976. K voprosu o sootnošenii upotrebljaemosti possessivnyx konstrukcij v sovremennyx slavjanskix jazykax [The relative frequency of use of possessive constructions in different contemporary Slavonic languages]. *Voprosy filologii* 5. 3–10. (Izdatel'stvo Leningradskogo universiteta).
- Johnson, Todd M. & Brian J. Grim. 2013. *The world's religions in figures: An introduction to international religious demography*. Chichester: John Wiley & Sons.

- Jung, Carl Gustav. 1921. *Psychologische typen*. Zurich: Rascher.
- Kibrik, Aleksandr E. 1995. Direct-oblique agreement of attributes in Daghestanian. In Frans Plank (ed.), *Double case: Agreement by Suffixaufnahme*, 216–229. New York: Oxford University Press.
- Koptjevskaja-Tamm, Maria. 2003. Possessive noun phrases in the languages of Europe. In Frans Plank (ed.), *Noun phrase structure in the languages of Europe*, 621–722. Berlin: Mouton de Gruyter.
- Krasovitsky, Alexander, Dunstan Brown, Greville G. Corbett, Matthew Baerman, Alison Long & Harley Quilliam. 2009. *Surrey database of short term morphosyntactic change: Case assignment on direct objects of negated transitive verbs*. Surrey: University of Surrey. <http://dx.doi.org/10.15126/SMG.17/3>.
- Kwon, Nahyun. 2017. Total reduplication in Japanese ideophones: An exercise in localized canonical typology. *Glossa: A Journal of General Linguistics* 2(1), 40. 1–31.
- Kwon, Nahyun & Erich R. Round. 2015. Phonaesthemes in morphological theory. *Morphology* 25. 1–27.
- Ladefoged, Peter & Ian Maddieson. 1990. Vowels of the world's languages. *Journal of Phonetics* 18. 93–122.
- Lakatos, Imre. 1976. *Proofs and refutations*. Cambridge: Cambridge University Press.
- Lander, Yury A. 2009. Varieties of genitive. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 581–592. Oxford: Oxford University Press.
- Lazarsfeld, Paul F. 1937. Some remarks on the typological procedures in social research. *Zeitschrift für Sozialforschung* 6. 119–139.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Lienhard, John H. 2000. *The engines of our ingenuity: An engineer looks at technology and culture*. Oxford: Oxford University Press.
- Matthews, P. H. 1997. *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.
- Mel'čuk, Igor'. 2006. Calculus of possibilities as a technique in linguistic typology. In Felix Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 171–205. Berlin: Mouton de Gruyter.
- Menzel, Thomas. 2018. Kanonische morphologische Komplexität im Sorbischen. *Lětopis Sorbisches Institut e.V. Bautzen* 65(2). 81–106.
- Michael, Lev. 2014. The Nanti reality status system: Implications for the typological validity of the realis/irrealis contrast. *Linguistic Typology* 18. 251–288.
- Murdock, George P. 1967. *Ethnographic atlas*. Pittsburgh: University of Pittsburgh Press.
- Mustajoki, Arto & Hannes Heino. 1991. *Case selection for the direct object in Russian negative clauses: II: Report on a statistical analysis (Slavica Helsinkiensia 9)*. Helsinki: Department of Slavonic Languages, University of Helsinki.
- Nettle, Daniel. 2018. *Hanging on to the edges. Essays on science, society, and the academic life*. Cambridge, UK: Open Book Publishers. <https://doi.org/10.11647/OBP.0155>.
- Newmeyer, Frederick J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 2010. On comparative concepts and descriptive categories: A reply to Haspelmath. *Language* 86. 688–695.
- Nichols, Johanna. 2019. Canonical tough cases. In Matthew Baerman, Oliver Bond & Andrew Hippisley (eds.), *Morphological perspectives*, 148–168. Edinburgh: Edinburgh University Press.

- Nikolaeva, Irina. 2013. Unpacking finiteness. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 99–122. Oxford: Oxford University Press.
- Nikolaeva, Irina & Andrew Spencer. 2013. Possession and modification – a perspective from canonical typology. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 207–238. Oxford: Oxford University Press.
- Parker, Steve. 2008. Sound level protrusions as physical correlates of sonority. *Journal of Phonetics* 36. 55–90.
- Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch & Barbara B. Lloyd (eds.), *Cognition and categorization*, 27–48. Hillsdale: Lawrence Erlbaum Associates.
- Round, Erich. R. 2017. Review of Gordon, Matthew K. Phonological typology. 2016. Oxford: Oxford University Press. *Folia Linguistica* 51. 745–755.
- Round, Erich. R. & Greville G. Corbett. 2017. The theory of feature systems: One feature versus two for Kayardild tense-aspect-mood. *Morphology* 27. 21–75. <https://link.springer.com/article/10.1007/s11525-016-9294-3> [open access].
- Rudner, Richard S. 1966. *Philosophy of social science*. Englewood Cliffs, N.J.: Prentice-Hall.
- Simpson, Adrian P. 1999. Fundamental problems in comparative phonetics and phonology: Does UPSID help to solve them. *Proceedings of the 14th international congress of phonetic sciences*, 349–352. Berkeley: University of California.
- Spencer, Andrew. 2013. *Lexical relatedness: A paradigm-based model*. Oxford: Oxford University Press.
- Spencer, Andrew. 2019. Canonical compounds. In Matthew Baerman, Oliver Bond & Andrew Hippisley (eds.), *Morphological perspectives*, 31–64. Edinburgh: Edinburgh University Press.
- Stanton, Elizabeth A. 2007. *The human development index: A history. (Working Paper Series No. 127)*. Amherst: Political Economy of Research Institute, University of Massachusetts.
- Stevens, Stanley S. 1951. Mathematics, measurement, and psychophysics. *Handbook of experimental psychology*, 1–49. New York: Wiley.
- Stump, Gregory. 2016. *Inflectional Paradigms: Content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.
- Stump, Gregory. 2017. The nature and dimensions of complexity in morphology. *Annual Review of Linguistics* 3. 65–83.
- Tabain, Marija, David Bradley & Defen Yu. 2019. Central Lisu. *Journal of the International Phonetic Association* 49. 129–147.
- Thornton, Anna. 2019. Overabundance in morphology. In *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acrefore/9780199384655.013.554>.
- Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.
- van den Berg, Helma. 2005. The East Caucasian language family. *Lingua* 115. 147–190.
- van Helden, W. Andries. 1993. *Case and gender: Concept formation between morphology and syntax (Studies in Slavic and General Linguistics 20)*. Amsterdam: Rodopi.
- Weber, Max. 1949 [1904]. “Objectivity” in social science and social policy. In Edward A. Shils & Henry A. Finch (eds. & trans.), *The methodology of the social sciences*, 50–112. Glencoe, Illinois: Free Press.
- Winch, Robert F. 1947. Heuristic and empirical typologies: A job for factor analysis. *American Sociological Review* 12. 68–75.
- Wisniak, Jaime. 2005. Development of the concept of absolute zero temperature. *Educación Química* 16. 104–113.

- Zaliznjak, Andrej A. 1973. O ponimanii termina 'padež' v lingvističeskix opisanijax [Interpreting the term 'case' in linguistic descriptions]. In Andrej A. Zaliznjak (ed.), *Problemy grammatičeskogo modelirovanija* [*Problems of grammatical modelling*], 53–87. Moscow: Nauka. [Reprinted in Andrej A. Zaliznjak. 2002. *Russkoe imennoe slovoizmenenie: s priloženiem izbrannix rabot po sovremennomu russkomu jazyku i obščemu jazykoznaniju* [Russian nominal inflection: With a supplement of selected works on contemporary Russian and general linguistics], 613–647. Moscow: Jazyki slavjanskoj kul'tury].
- Zaliznjak, Andrej A. 1977. *Grammatičeskij slovar' russkogo jazyka: slovoizmenenie* [*A grammatical dictionary of Russian: Inflection*]. Moscow: Russkij jazyk. [A fourth, corrected edition appeared in 2003, Moscow: Russkie slovari].