On the relationship between typology and the description of Uralic languages

Miestamo, Matti Markko Petteri

2018

unspecified
publishedVersion

ON THE RELATIONSHIP BETWEEN TYPOLOGY AND THE DESCRIPTION OF URALIC LANGUAGES

Matti Miestamo
University of Helsinki

Abstract. Language typologists are dependent on data provided by descriptive linguists working on individual languages, who, in turn, benefit from typologists' results, which give them new insights into the properties of their respective languages. The article addresses the relationship between typology and Uralic descriptive linguistics. The state of description of Uralic languages is examined by surveying the availability of descriptive sources on Uralic languages in the Glottolog database, which is widely used by typologists. Ways in which Uralic studies and language typology can be brought closer to each other for the benefit of both fields are discussed giving a number of recommendations for writing typologically-oriented grammars of (Uralic) languages. Finally, the use of typologically informed questionnaires in language description is briefly addressed.

1. Introduction

The organizers of the workshop "Typology of Uralic languages: towards better comparability" at the SLE conference in Naples in 2016 defined their goals as follows: "The main goal of the workshop is to increase the level of comparability of the Uralic languages, and to promote the integration of Uralic studies into contemporary linguistics." In accordance with these goals, this paper presents a typologist's perspective on the state of documentation and description of Uralic languages, focusing on the latter, and discusses ways in which Uralic studies and language typology can be brought closer to each other for the benefit of both fields.

Language typology and language description live in a symbiotic relationship. On the one hand, descriptive work provides the empirical foundations for typology – without the work of descriptive linguists

studying individual languages typological comparison would not be possible. On the other hand, typology feeds description with theory and ideas and provides descriptive linguists with better theoretical tools for understanding the phenomena they encounter in the languages they study. In other words, language typology, and thus also general linguistics more broadly, is dependent on data provided by researchers focusing on particular languages, who then, in turn, benefit from typological knowledge when trying to understand the phenomena they encounter in their respective languages of study. Typology and descriptive linguistics share the common goal of unraveling, describing and understanding the world's linguistic diversity, and *Diversity Linguistics* has been suggested as name for the subfield of linguistics that documentary and descriptive work and comparative (typological and historical) linguistics constitute[1].

In this definition, descriptive and comparative studies of Uralic languages belong to the field of Diversity Linguistics. So let us ask the question how visible and well-integrated Uralic linguistics is in this field and in typological linguistics in particular. On the one hand, we may ask whether, from the point of view of Uralistics, one can be satisfied with the way Uralic languages are represented in typological studies. On the other hand, from a different point of view, we may ask whether one can be satisfied with how typology is taken into account in Uralistics. A lot can be learned about the situation by looking at the descriptive work done on Uralic languages and paying attention to the extent to which Uralic descriptive work is available to general linguistic audiences and to the extent to which Uralic descriptive work has been written in a typological perspective. In Section 2, I will make a survey of Uralic descriptive work as represented in the Glottolog database (glottolog.org).

Before going into that survey, a few words on the relationship between documentation and description are in order. There is some variation as to how different researchers understand these concepts and use these terms. Explicit definitions have been given by Himmelmann (1998) and Lehmann (2001). The latter is quoted here:

> **Documentation** of a language is an activity (and, derivatively, its result) that gathers, processes and exhibits a sample of data of the language that is representative of its linguistic structure and gives a fair impression of how and for what purposes the language is used. Its aim is to represent

---

1   http://dlc.hypotheses.org/1

the language for those who do not have access to the language itself. **Description** of a language is an activity (and, derivatively, its result) that formulates, in the most general way possible, the patterns underlying the linguistic data. Its aim is to make the user of the description understand the way the language works (Lehmann 2001).

Lehmann thus makes an explicit distinction between description and documentation, but sometimes both activities are lumped together under labels such as language documentation or descriptive linguistics. Himmelmann (1998) and many others have argued for the status of documentary linguistics as an independent subfield of linguistics, covering documentation in the narrow sense but not description. As also Himmelmann (2016) points out, documentation and description presuppose each other – meaningful presentation of data requires descriptive analysis and description requires data – and in practice these two activities are often done simultaneously. In this paper, I will use the terms in the narrow sense as defined by Lehmann. And since it is descriptive work that is most relevant to typology, I will mostly only talk about description an descriptive linguistics, but it should be remembered that simultaneous or earlier documentary work is always involved in a descriptivist's work.

The Uralic language family is one of the best studied families in the world. There is a long tradition of fieldwork starting from the early 19th century with Anders Sjögren and Mathias Alexander Castrén, see e.g. Hovdhaugen et al. (2001) for an overview and discussion. This 200-year long tradition has produced a lot of documentary and descriptive material on Uralic languages and an extensive literature on the history and genealogical classification of the family. However, at least until very recently, the main focus in Uralic studies has remained historical-comparative, rather than typologically oriented, and the Uralic family does not necessarily appear very well-described to linguists working outside the Uralic tradition. The documentary and descriptive materials produced are typically text collections, dictionaries, chrestomathies, or even more extensive grammars, but, as also noted by the workshop organizers, descriptive grammars written in a modern typological-functional perspective are largely lacking to date.

What I mean by this are grammars that can be roughly characterized by the following qualities: they are informed by up-to-date linguistic theory, especially by typological research, they are organized (also) according to linguistic functions rather than (only) by the formal categories of the language in question and are thus more readily usable

for typological-comparative research, and consequently they pay much more attention to syntax and various semantic and pragmatic issues than more traditional grammars, which often do not go very much, if at all, beyond phonology and morphology. I will come back to the desired properties of a good contemporary grammatical description further below, but this should suffice as background for the following survey.

2. Survey of the state of description of Uralic languages

In order to get a more concrete view of how the state of description of the Uralic family might appear to a non-Uralicist typologist, I made a survey of the sources listed for Uralic languages in the Glottolog database (glottolog.org). Glottolog is a comprehensive and up-to-date catalogue of the world's languages, language families and dialects. Most importantly for the present purposes, it features a comprehensive collection of bibliographical data for the world's languages, especially for lesser-known ones, containing roughly 315,000 entries as of June 2017. Today, it is the first place to go for most typologists when looking for sources on languages to be included in their comparative work. The references in the database are classified into 16 different document types: Bibliographically oriented, Comparative-historical study, Dialectologically oriented, Dictionary, Ethnographic work, Grammar, Grammar sketch, Handbook/overview, New Testament, Phonology, Sociolinguistically oriented, Some very small amount of data/information on a language, Text, (Typological) Study of a specific feature, Unknown, Wordlist. I will not go into the definition of each of these types here, but focus on the two most relevant ones for the present survey, i.e. on the most typical sources that typologists use when comparing languages: grammars and grammar sketches. A grammar is an extensive description of most elements of the language's grammar, typically 150 pages and beyond, whereas a grammar sketch is a less extensive description of many elements of the grammar, typically around 50 pages.

Let us now take a look at what a typologist interested in finding sources on Uralic languages will find in these two categories in the Glottolog database. The database (as of June 2017) features 48 Uralic languages and lists altogether 1,212 references for them. 158 of these entries are classified as grammars and 200 as grammar sketches. All these numbers should be taken with a grain of salt: firstly, some existing grammars and grammar sketches may be missing from the database (e.g.

Grünthal 2015), and secondly, some references are not classified into any type and some of them are classified automatically, so there will be errors. However, the two document types that I am most interested in here are relatively likely to be easily recognizable as compared to some other document types, so it is less likely that the entries belonging to them would be misclassified or unclassified. It is especially unlikely that the kind of extensive modern typologically-oriented grammatical descriptions that we are most interested here, would be misclassified as something else than a grammar. A proper bibliographic analysis made by an expert in Uralic linguistics would very probably yield a somewhat different result, but the picture given by a survey of this general database is certainly suggestive of the bibliographic reality. And most importantly, it is exactly this picture that we are interested in here when we are trying to find out how the Uralic sources offer themselves to a non-Uralicist typologist.

Taking a closer look at individual languages and the grammars and grammar sketches available for them, we may first note that out of the 48 Uralic languages in the database, there are three that do not have either a grammar or a grammar sketch, 14 that have only a grammar sketch but no grammar and 31 that have a grammar, see Table 1. Note that the names of the languages and the classification into genealogical groups are taken as such from the Glottolog database and do not reflect my position on how the languages should be named or classified.

Table 1. Uralic languages with information on existence of grammar, grammar sketch only or neither in the Glottolog database (as of June 2017)

Neither grammar nor grammar sketch	
Finnic	Ludian
Saami	Inari Saami, Kemi Saami

Grammar sketch but no grammar	
Finnic	Ingrian, Livvi, Kven Finnish, Võro
Saami	Ume Saami
Permian	Komi-Yazva
Khantyic	Surgut Khanty, Southern Khanty
Mansic	Central Mansi, Southern Mansi
Samoyedic	Tundra Enets, Forest Nenets, Kamas-Koibal, Mator-Taigi-Karagas

Table 1. Continuation

GRAMMAR	
Finnic	Estonian, Karelian, Liv, Finnish, Tornedalen Finnish, Veps, Votic
Saami	Skolt Saami, Akkala Saami, Kildin Saami, Ter Saami, Lule Saami, Northern Saami, Pite Saami, Southern Saami
Mordvin	Erzya, Moksha
Mari	Eastern Mari, Western Mari
Permian	Komi-Permyak, Komi-Zyrian, Udmurt
Khantyic	Vach-Vasjugan, Kazym-Berezov-Shuryshkar Khanty, Obdorsk Khanty
Mansic	Northern Mansi
Hungarian	Hungarian
Samoyedic	Tundra Nenets, Forest Enets, Selkup, Nganasan

Looking at the sheer numbers, the state of description of the Uralic family looks relatively good: 65 % of the languages have a grammar and only 6 % lack both grammar and grammar sketch[2]. Furthermore, languages with grammars are found in every major branch of the family. One could of course argue that for a relatively small family with such a long research tradition, even these numbers are not extremely good. Be it as it may, for a more accurate picture, we need to examine what kinds of entries we find in the category of grammar in the 31 languages. Table 2 shows the numbers of items classified as grammars by publication year. Note that there are a few items for which the year of publication is not coded in the database in such a way that it would come up in the search, so that the total number is a bit lower than the number given above.

The pace of publication is rather slow until the 1960s when suddenly the volume doubles within one decade – this increase is largely due to the appearance of many Russian-language grammatical descriptions of Uralic languages spoken in the Soviet Union. There is no denying that even the oldest grammars from the 19th century contain a lot of interesting and useful material for today's researchers, but for the present purposes, i.e. to find out to what extent modern typologically informed

---

2   One of these three, namely Kemi Sami, is long extinct, so the existence of a grammar or grammar sketch can hardly be expected.

Table 2. Numbers of Uralic grammars by year in the Glottolog

year	number of grammars
1800–1899	8
1900–1959	16
1960–1969	32
1970–1979	12
1980–1989	19
1990–1999	31
2000–2009	19
2010–2017	12
total	**149**

grammars exist for Uralic languages, it is better to focus on the more recent periods in this survey. It is beyond the scope of this paper to go through each publication classified as grammar in the Glottolog in detail and try to determine what the most comprehensive and up-to-date description for each language is. What I have done instead is to try to get a picture of how many Uralic languages possess a grammar that could possibly be characterized as a comprehensive, extensive, modern, typologically-informed grammatical description. There are no hard-and-fast criteria for this category, but to get some idea, I have started with rather mechanical criteria and listed the publications on Uralic languages classified as grammars in the Glottolog published within the last 25 years, i.e. after the year 1992, and that are more than 250 pages in length. The logic behind these criteria is that a grammar that is shorter than 250 pages or published more than 25 years ago is not a very likely candidate for the kind of description we are interested in here. I am fully aware of the possibility that these arbitrary cut-off points may do unjustice to some grammars that fail the criteria by a close margin. The references satisfying these criteria are listed in Table 3. To keep the table from growing too extensive, I have excluded the references for the three major national languages Estonian, Finnish and Hungarian, since these languages can be considered (relatively) well-described.

Table 3. Uralic publications, classified as grammars in the Glottolog, with more than 250 pages and published since 1992 (excluding Estonian, Finnish and Hungarian)

LANGUAGE	AUTHORS	YEAR	TITLE	PP
Votic	Markus & Rožanskij	2011	Sovremennyj Vodskij Jazyk: Tektsy i Grammatičeskij Očerk	720
	Winkler	1997	Krewinisch: zur Erschliessung einer ausgestorbenen ostseefinnischen Sprachform	468
Liv	Winkler	1994	Salis-Livische Sprachmaterialien	443
Veps	Brodskij	2008	Samoučitel' Vepsskogo Jazyka	428
Eastern Khanty	Filchenko	2010	Aspects of the Grammar of Eastern Khanty	555
	Filchenko	2007	A Grammar of Eastern Khanty	614
Obdorsk Khanty	Nikolaeva	1995	Obdorskij dialekt Chantyjskogo jazyka	262
Eastern Mari	Alhoniemi	1993	Grammatik des Tscheremissischen (Mari)	253
Western Mari	Alhoniemi	1993	Grammatik des Tscheremissischen (Mari)	253
	Vasil'ev & Učaev	2003	Marijsko-russkij slovar': [dlja izucajuscich marijskij jazyk]	293
Komi-Permyak	Fedjunëva	1998	Komi jazyk: enciklopedija	607
Komi-Zyrian	Avril	2006	Parlons komi	261
Udmurt	Suihkonen	1995	Udmurt – English – Finnish dictionary with a basic grammar of Udmurt	326
	Kel'makov & Hännikäinen	1999	Udmurtin kielioppia ja harjoituksia	319
Skolt Sami	Feist	2010	A Grammar of Skolt Saami	478
Northern Sami	Nickel	1994	Samisk grammatikk	539
	Nickel & Sammallahti	2011	Nordsamisk grammatikk	691
	Fernandez	2000	Parlons Lapon: Les Sames	347

LANGUAGE	AUTHORS	YEAR	TITLE	PP
	Pope & Sárá	2004	Eatnigiella: giellaoahpu váldogirji	293
Pite Sami	Wilbur	2013	A corpus-based grammar of spoken Pite Saami	316
	Wilbur	2014	A grammar of Pite Saami	314
Forest Enets	Siegl	2013	Materials on Forest Enets, an indigenous language of Northern Siberia	523
Tundra Nenets	Nikolaeva	2014	A Grammar of Tundra Nenets	526

It should first be noted that one publication is listed twice, namely Alhoniemi 1993 for both Eastern and Western Mari, and a couple of items are quasi-duplicates, namely Filchenko 2007/2010 and Wilbur 2013/2014. Counting these only once, we end up with 20 separate grammars for 15 Uralic languages satisfying the age and length criteria. It is, however, clear that many of these publications are not the kind of modern comprehensive descriptive grammars that we are interested in here: several of them are pedagogical textbooks rather than descriptive grammars (Avril 2006, Kel'makov and Hännikäinen 1999, Fernandez 2000 and Pope and Sárá 2004), a couple of them are primarily dictionaries and the grammar part is not very extensive (Vasil'ev and Učaev 2003, and Suihkonen 1995), and some are philological studies of textual material rather than descriptive grammars (Winkler 1994, 1997). Leaving these publications out of consideration, we are left with 12 grammars for 12 languages. As I said above, it is not possible to determine categorically, which publications can be seen as modern typologically informed grammatical descriptions, and defining such a clear line is beside the point. In my somewhat subjective evaluation, the following seven are the clearest candidates: Markus and Rožanskij 2011, Filchenko 2010, Feist 2010[3], Nickel and Sammallahti 2011, Wilbur 2014, Siegl 2013 and Nikolaeva 2014. Adding the three well-described major national languages, the following 10 Uralic languages appear as best described from a typologist's perspective: Estonian, Finnish, Votic, Eastern Khanty, Hungarian, Skolt Saami, Northern Saami, Pite Saami, Forest Enets and Tundra Nenets. Needless to say, this does not mean

---

3   Note that a revised version of Feist 2010 has been published as Feist 2015, but it is not yet in the Glottolog database.

that all these languages would be fully or even satisfactorily described, but we can confidently conclude that in terms of the sources available for typological comparison, they stand out from the rest of the Uralic family.

The picture given by the survey of Uralic descriptive materials in the Glottolog suggests that overall the state of description of the Uralic family is not very good as regards grammars written in a modern typologically oriented perspective. To date, most Uralic languages lack a comprehensive and extensive grammar informed by modern typological theory. This is quite surprising given the long tradition of research on Uralic languages. Looking at language description in a word-wide perspective, within the last two decades, we see a lot of modern typologically informed work done on languages belonging to families with little or no previous research tradition in remote parts of the world. In the lack of a specialized research tradition and specialized community of researchers, such languages are often studied by general linguists who come to the field with a typologically-oriented theoretical mindset, and they end up being described according to the most up-to-date methodological and theoretical desiderata. In a way, Uralic languages have been a victim of the strong historical-comparative tradition that has not prioritized synchronic typologically-oriented work. Certainly this does not mean that the Uralic family would be less thoroughly documented and described overall than a group of languages for which such modern descriptions exist but on which altogether only one or a few linguists have ever worked. The long Uralic tradition puts Uralic linguists in a very good position to extend the volume and raise the level of description of Uralic languages and disseminate this information to the general linguistic community.

Luckily, more and more efforts towards this end are being made. In the survey of sources in the Glottolog database, there was a marked difference between the latest decade and earlier times. Currently several projects documenting and describing Uralic languages are on the way and the interaction between typologists and Uralicists is more lively than ever before. One major event triggering mutual awareness and cooperation was the Uralic Typological Database Conference in Vienna in September 2008 that brought together typologists and Uralicists working on shared topics. Although the main idea of that conference was to discuss and launch a project that would result in an extensive database of typological information on Uralic languages, many typologists voiced the opinion that from a typologist's perspective, a more

interesting prospect would be the production of good quality grammars. Throughout this paper, I have emphasized the need of grammatical descriptions for the typological and general linguistic community. It is the perspective of this paper, but it should not be forgotten that grammars are not written for the typologist only and not even primarily – describing and analysing Uralic languages in a typologically informed perspective will advance our understanding of Uralic languages and benefit first and foremost Uralic linguistics itself.

3. On writing grammars

What, then, should a descriptive linguist working on a Uralic language (or any other language for that matter) take into account in writing a grammar so that it would serve specialists as well as typologists and other general linguists in the best possible way. I argue that a good grammar should be: typologically informed, framework-free, accessible, functionally oriented, multidimensional and comprehensive. This section will address these six points one by one.

The first point has been mentioned several times already in this paper. A good grammatical description is typologically informed. Typology provides important theoretical background for understanding language-particular phenomena. Knowledge about other languages helps the descriptive linguist to see her/his data in a broader cross-linguistic context. On the one hand, typology provides a broader perspective for understanding the data, helps to ask relevant questions about the data, suggests hypotheses and possible analyses, sees connections between different phenomena and fills gaps that might otherwise go unnoticed. On the other hand, typological literature provides a meta-language for description (more on metalanguage further below). Publications specifically intended to provide a broad typological-theoretical background for language description include Payne (1997) and Shopen (2007), and further useful general works include Aikhenvald and Dixon (2017), Dixon (2009–2012), Dryer and Haspelmath (2013), Haspelmath et al. (2001), and Song (2011); naturally any typological literature on a specific domain of grammar is relevant for the description of that domain in a given language. Another concrete way to integrate typological knowledge in fieldwork and grammar writing is to make use of various typologically oriented questionnaires, which I will come back to at the end of this section. For more discussion on the relevance of

typology for language documentation and description, see Bond (2010), Epps (2011), Zúñiga (2012), Sandman (2013) and Himmelmann (2016).

The second requirement on my list is that a good grammatical description should be framework-free. This rests on the old structuralist principle that each language should be described in its own terms and one should be aware of the dangers of letting the categories of other languages influence the way a language is described. Recently, Haspelmath (2010a) has emphasized this point, introducing the term framework-free grammatical description. It is not only the model of other languages that bear the danger of imposing alien categories on the language that is being described, but also grammatical theories and frameworks that are based on the model of languages that have received most attention in linguistic theory, most notably English and other major European languages[4]. This connects to recent discussions on the nature of linguistic categories (see Lazard 1992, Dryer 1997, Haspelmath 2007, 2010b, Miestamo 2013)[5]. Haspelmath (2007) makes a distinction between categorial universalism and categorial particularism. In a universalist view, there are universal cross-linguistic categories and the categories of particular languages are language-particular realizations of these – this view is held e.g. by many adherents to the Chomskyan generative paradigm. In a particularist view, grammatical categories are particular to each individual language, cross-linguistic categories do not exist, and the similarities between categories observed in two or more languages are due to similarities on the functional level (semantics and pragmatics) – this view was prominent e.g. in early American structuralism (see Joos 1957) and is today held by many typologists and descriptive linguists. The framework-free approach to language description is compatible with categorial particularism. Essential in this approach is to find out and describe the properties of the linguistic forms and combinations of forms, paying attention to their function and distribution, rather than to look for categories assumed a priori on the basis of categories present in other languages or established frameworks. It may appear somewhat controversial to promote typological knowledge as a guide to understanding the properties of a language and at the same time warn against the influence of knowledge of other languages on the analysis and description. The key lies in working bottom-up and being true to the data and giving typology the role of guide and companion, but not

---

4   Gil (2011) characterizes fieldwork as a process of unlearning, i.e. learning to get rid of the a priori categories provided by Euro-centric linguistic frameworks.
5   See also the discussion on categories in *Linguistic Typology* 20:2, 2016.

that of a model. Haspelmath (2016) stresses the importance of a clear separation between the comparative concepts used in typology and the language-particular descriptive categories used in language description. It is perhaps worth adding that the descriptive categories established by a descriptive linguist are an abstraction the linguist makes from the language used by the language community. The better these descriptive categories correspond to that reality, i.e. the better they reflect the psycho-social reality of the speakers, the more succesful and adequate we can estimate the description to be.

The next point on the good grammars wish list is accessibility. In a nutshell this means that the grammar uses a theoretical background and a metalanguage/terminology that is shared by or at least understandable by as many linguists (and other readers) as possible. Many writers, most notably Dryer (2006), have made the point that grammars should not be written using the metalanguage of a theory that is not widely shared and established and that might go out of fashion and be forgotten in the relatively near future. There are examples of grammars written within such frameworks, whose usability for today's linguistic community is limited, e.g. the tagmemic grammars of Bolivian languages published in Matteson 1967 or the transformational grammar of Hidatsa by Matthews (1965). Basic Linguistic Theory (Dixon 1997, Dryer 2006) has been proposed as a term referring to the common theoretical background shared by most linguists. BLT goes ultimately back to traditional grammar and has accumulatively integrated theoretical concepts and terms from different approaches over the years – structuralism, generative grammar, and in the last few decades especially from typology. Theoretical innovations that have proved to be of lasting value have become part of BLT, but concepts that have been abandoned as linguistic theory has progressed have not. This is the original idea behind the concept of BLT. More recently, Dixon has taken a different approach and introduced BLT as a more explicitly defined framework (Dixon 2009–2012). This clearly departs from the original idea. Haspelmath (2010a) has warned against BLT functioning as a framework that has the same dangers of imposing a priori categories as any other framework. In my view, BLT understood in its original sense as the commonly shared theoretical background providing an accessible metalanguage for description, provides a viable perspective for grammatical description.

A few more words on metalanguage are in order. In the particularist view, categories are language-specific and consequently, terminological similarity between descriptions of different languages is not

a theoretical necessity. In principle, the language-particular categories could be referred to with arbitrary numbers and letters[6], which would indeed emphasize the language-particularity of categories in that there is no terminology to suggest connections between categories across languages. However, such a grammar is very difficult to use as terminology does not give any hints as to the functions of the categories and the comparability between such grammars is very low. A better approach, as also recommended by Haspelmath (2010b), is to use similar terminology for similar categories across languages: while categories are language-particular and no cross-linguistic categories are recognized, we can still observe similarities between categories in different languages and it is useful to refer to them with similar terms, as long as we remember that terminological similarity does not mean sameness of categories.

More technical points related to accessibility have to do with glossing and transcription. These are especially important when writing grammars for wider audiences beyond the Uralicist circles. The Leipzig Glossing Rules[7] have become a common standard shared and accepted by a wide community of linguists. They are designed to be applicable across languages and provide a clear and concise way of glossing and presenting examples of languages unknown to the reader. The Finno-Ugric phonetic transcription system is not widely known and using IPA instead will enhance the accessibility of the grammar for wider audiences.

Next on the list we have the requirement that a good grammatical description should be functionally oriented. This means that it is organized (also) according to function rather than (only) according to form. A formally organized grammar is divided into chapters on the basis of the formal classes identified in the language, and typically contains chapters discussing the forms and paradigms of nouns, verbs, adjectives, adverbs etc, and possibly also noun phrases and clause structure. A functionally oriented grammar builds its chapters around functional domains such as argument marking, possession, tense-aspect-mood-evidentiality, negation, interrogatives and commands etc, and organizes them according to which functional domains are grammaticalized in the language. No grammar can be exclusively functional in its orien-

---

6   This is not very far from the reality of some tagmemic grammars, e.g. those in Matteson (1967); in tagmemics the practice to describe each language in its own terms was made explicit in this way.

7   https://www.eva.mpg.de/lingua/resources/glossing-rules.php

tation – even in a strongly functionally-oriented grammar, the formal classes and their inflectional properties have to be presented, but then their meanings and functions are introduced and discussed in later chapters that are organized on functional grounds. A good example of a functionally-organized grammar is Frajzyngier's (2002) grammar of Hdi; see Frajzyngier and Shay (2003) for more on the role of functional domains in grammatical description.

What are the benefits of functional organization? A core question is who the grammar is written for. A formally-based grammar may serve specialists of the language and the language family but be less readily usable for typologists and other general linguists. A Uralicist opening a grammar of, say, Udmurt may be looking for information on the properties of a given form that has cognates in other Uralic languages and is well served by a formally oriented grammar. A typologist is more likely to be looking for information on how a particular function, e.g. negation or evidentiality, is expressed in Udmurt and is thereby much better served by a functional orientation. From the point of view of serving a readership as wide as possible, a good approach is to organize the grammar in terms of function as far as possible, and then provide a good index of forms referring to the pages on which each form is treated. When grammars are written in a typological perspective as described above, a functional orientation presents itself quite naturally. Typological comparison between languages is primarily based on function rather than form (see Stassen 1985, Haspelmath 2010b, Miestamo 2013), and asking typologically relevant questions about the grammar of a language is very often about how the language expresses a given meaning/function.

Next on the list we have comprehensiveness. This has been alluded to already in the discussion in the previous sections, and simply means that a description should cover all domains of the grammar of the language so as to serve the needs of readers interested in any aspects of the grammar of the language. Attention should be paid comprehensively to different functions and functional domains suggested in typological literature. The focus and organization of the grammar should then follow those functional domains that are found to be grammaticalized in the language under description, and what is special and salient in the grammar of the language should be given proper attention in the grammar.

This leads further to the question what the limits of grammar are – what belongs to a grammar and what should be treated elsewhere? One

such limit is between grammar and lexicon, more concretely, what is the division of labour between a grammar and a dictionary. But there are other dimensions too, and we are brought to the last point on our list of requirements: multidimensionality. This means that the grammar should not stay within the limits of the strictly linguistic, but should provide information on different aspects of the ecology of the language that would shed light on and deepen our understanding of the grammatical properties of the language. As has been emphasized in ecological approaches to language (e.g. Mühlhäusler 1996), languages are in close interaction with their physical, social and cultural environments and in order to really understand what is going on in a language, we need information on these ecological aspects of the language. Grammars should ideally give information on the language community's history, ethnography, cultural practices, social structure, geographical and biological surroundings, contacts with neighbouring groups etc. This idea of multidimensionality has recently been developed within the Helsinki Area and Language Studies (HALS) research community, and we have termed this approach *rich description*, see Nichols (2017) for examples. Uralic linguistics is in fact in a very good position to promote such an approach, since the tradition started in the early 19th century has always included a very strong ethnographic component, and a lot of material has been collected throughout the history of Uralic studies that is directly relevant to enriching grammatical descriptions with various dimensions of the ecology of the languages. A central idea with rich description is that the ecological aspects are not merely given as background in an introductory chapter, but they are truly integrated into the description at points where they are relevant to the grammatical phenomenon under discussion. Modern digital technology provides new opportunities for integrating ecological dimensions into the grammar. Grammars can include links to more detailed descriptions of the relevant aspects of language ecology (research publications, videos, text materials etc). Similarly grammars can be directly linked to dictionaries and corpora which also contribute to the comprehensive and multidimensional nature of the grammar.

I have now discussed the main points that I consider to be of central importance if grammatical description is to benefit from typological work and in turn be maximally useful for typologists and other general linguists. Before concluding the paper, I will say some words about a topic that was already briefly mentioned in connection with the requirement of typological informedness, namely the use of (typological) ques-

tionnaires in language description. A wide variety of questionnaires are available on the (still existing) web pages of the former department of linguistics at the MPI Leipzig[8]. Some of these are originally intended for typological comparison and some are meant for grammatical description, but even the latter are very often designed in a typological perspective so that the questions to be asked in grammatical description stem from typological knowledge. The following quote from the website's description expresses the purpose succinctly:

> This website contains tools for use in field linguistics and language description. Most of the items on the website are questionnaires designed to assist in eliciting data in such a fashion that the data will be comparable across languages. The idea behind this website was that field linguistics should be typologically informed and that the results of field work should be of typological interest. The questionnaires and other tools presented here help the field linguist understand what questions might be of typological (and theoretical) interest and guide the linguist in both eliciting data and extracting information from naturalistic texts.

In addition to the questionnaires, the website provides other useful information for language description, such as references to useful literature, and it also hosts the Leipzig Glossing Rules. Unfortunately, as the Leipzig department has been closed down, the future of the website is uncertain. A new initiative of archiving and providing access to questionnaires and stimulus kits is being developed at the CNRS Fédération de Recherche, Typologie et Universaux Linguistiques[9].

Different types of questionnaires can be identified along the following dimensions. They may show differences in scope: some questionnaires aim at a full description of a language (e.g. the Lingua descriptive studies questionnaire, Comrie and Smith 1977), some others target a specific domain of grammar (e.g., the negation questionnaire by Miestamo 2016 or the tense-aspect questionnaire in Dahl 1985), and finally some questionnaires are interested in eliciting data to address a more narrow typologically interesting question. Questionnaires may also be different with respect to the targeted user group: some questionnaires are meant for the linguist as a sort of checklist on what aspects of grammar to pay attention to in the grammatical description (e.g. Comrie

8   https://www.eva.mpg.de/lingua/tools-at-lingboard/tools.php
9   http://tulquest.huma-num.fr

and Smith 1977 or Miestamo 2016), while others are meant to be used directly with a native speaker; these can be translation questionnaires that provide a sentence in the metalanguage to be translated to the target language (e.g. Dahl 1985), or they can be picture questionnaires or video stimuli (e.g. the Pear Stories). All these different types of questionnaires can be used as aids to make a grammatical description more typologically oriented and increase its comparability with other descriptions.

As a concrete example, we may take a brief look at the negation questionnaire mentioned above (Miestamo 2016). The questionnaire aims at a comprehensive coverage of the domain of negation. It is based on contemporary typological and theoretical knowledge of negation and addresses different aspects of negation identified as typologically and theoretically interesting in the literature. The main topics covered are organized under three sections: 1. clausal negation – standard negation, negation in non-declaratives, negation of stative predications, negation in dependent clauses; 2. non-clausal negation – negative replies, negation of indefinite pronouns, negative case, derivation and adpositions; 3. further aspects of negation – scope of negation, negative polarity, negation and case marking, reinforcing negation, negation and complex sentences. The questionnaire was originally created for the purposes of a volume on negation in Uralic languages (Miestamo et al. 2015). The first part of the volume consists of 17 chapters each describing one Uralic language on the basis of the questionnaire. This unified description maximizes comparability between Uralic languages, and given the typological basis of the questionnaire, it also makes the Uralic data easily accessible for typological audiences. The version of the questionnaire published in the volume has been somewhat adapted to the needs of the description of Uralic languages – examples are taken from Uralic languages and special attention is paid to phenomena that are known to be prominent in Uralic, e.g., negative verbs. The revised 2016 version is more general in its orientation. One of the main goals of the volume is to bring typology and Uralic linguistics closer together to the advantage of both fields and thereby enhance advances in the study of negation in both typology and Uralic linguistics. Furthermore, there is a direct benefit for language description as the chapters in the volume provide an excellent basis for chapters on negation in future grammars of the respective languages. Such typologically oriented thematic volumes focusing on other functional domains as well would be very useful for the description of Uralic languages in general.

4. Conclusion

In this paper I have surveyed the state of description of the Uralic language family and provided some ideas and advice for projects aiming at producing grammars of Uralic (and other) languages. I started by discussing the concepts of documenation and description, surveyed the bibliographic situation of Uralic as shown by the Glottolog database and concluded that most Uralic languages still lack a comprehensive and extensive grammar informed by modern typological theory. The last part of the chapter concentrated on how to remedy the situation and what to take into account in writing such grammars. The goal of this paper has not been to provide an exhaustive bibliographic analysis of Uralic grammatical descriptions nor to give a comprehensive guide for grammar writers starting from how to prepare for fieldwork. Rather I wanted to give a general picture of the state of description of the Uralic family and take up some points that I deem important for grammar writing from a typological perspective. I hope these thoughts can be useful for linguists working on the documentation and description of Uralic languages and generate more discussion on the current situation and on the best practices.

Address
Matti Miestamo
General Linguistics
P.O. Box 24 (Unioninkatu 40)
FI-00014 University of Helsinki
Finland
E-mail: matti.miestamo@helsinki.fi

References

Aikhenvald, Alexandra and Robert M. W. Dixon, eds. (2017) *The Cambridge handbook of linguistic typology*. Cambridge: Cambridge University Press.

Alhoniemi, Alho (1993) *Grammatik des Tscheremissischen (Mari)*. Hamburg: Buske.

Avril, Yves (2006) *Parlons komi*. (Collection Parlons.) Paris: L'Harmattan.

Bond, Oliver (2010) "Language documentation and language typology". *Language Documentation and Description* 7, 238–261.

Brodskij, Igor' V. (2008) *Samoučitel' Vepsskogo Jazyka*. Sankt Peterburg: Sanktpeterburgskoe Vepsskoe Obščestvo.

Comrie, Bernard, and Norval Smith (1977) "Lingua descriptive studies: questionnaire". *Lingua* 42:1, 1–72.

Dahl, Östen (1985) *Tense and aspect systems*. Oxford: Blackwell.

Dixon, Robert M. W. (1997) *The rise and fall of languages*. Cambridge: Cambridge University Press.

Dixon, Robert M. W. (2009–2012) *Basic linguistic theory.* 3 Vols. Oxford: Oxford University Press.

Dryer, Matthew S. (1997) "Are grammatical relations universal?". In Joan Bybee, John Haiman, and Sandra Thompson, eds. *Essays on language function and language type: dedicated to T. Givon*, 115–143. Amsterdam: Benjamins.

Dryer, Matthew S. (2006) "Descriptive theories, explanatory theories, and basic linguistic theory". In Felix Ameka, Alan Dench, and Nicholas Evans, eds. *Catching language: issues in grammar writing*, 207–234. Berlin: Mouton de Gruyter.

Dryer, Matthew S., and Martin Haspelmath, eds. (2013) *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/>. Accessed on 28.10.2018.

Epps, Patience (2011) "Linguistic Typology and Language Documentation". In Jae Jung Song, ed. *The Oxford handbook of linguistic typology*. Oxford: Oxford University Press.

Fedjunëva, Galina V. (1998) *Komi jazyk: ènciklopedija*. Moskva: Izdat. Dik

Feist, Timothy (2010) *A grammar of Skolt Saami.* Doctoral dissertation, University of Manchester.

Feist, Timothy (2015) *A grammar of Skolt Saami.* (Mémoires de la Société Finno-Ougrienne, 273.) Helsinki: Finno-Ugrian Society.

Fernandez, Jocelyne M. M. (2000) *Parlons lapon: les sames*. (Collection Parlons.) Paris: L'Harmattan.

Filchenko, Andrey Y. (2007) *A grammar of Eastern Khanty*. Doctoral dissertation. Houston: Rice University.

Filchenko, Andrey Y. (2010) *Aspects of the grammar of Eastern Khanty*. Tomsk: TSPU Press.

Frajzyngier, Zygmunt, with Erin Shay (2002) *A grammar of Hdi*. (Mouton Grammar Library 21.) Berlin: Mouton de Gruyter.

Frajzyngier, Zygmunt and Erin Shay (2003) *Explaining language structure through systems interaction.* (Typological Studies in Language, 55.). Amsterdam: Benjamins.

Gil, David (2001) "Escaping Eurocentrism: fieldwork as a process of unlearning". In P. Newman and M. Ratliff, eds. *Linguistic fieldwork*, 102–132. Cambridge: Cambridge University Press.

Grünthal, Riho (2015) *Vepsän kielioppi*. (Apuneuvoja suomalais-ugrilaisten kielten opintoja varten.) Helsinki: Suomalais-Ugrilainen Seura.

Haspelmath, Martin (2007) "Pre-established categories don't exist: consequences for language description and typology". *Linguistic Typology* 11:1, 119–132.

Haspelmath, Martin (2010a) "Framework-free grammatical theory". In Bernd Heine and Heiko Narrog, eds. *The Oxford handbook of grammatical analysis*, 341–365. Oxford: Oxford University Press.

Haspelmath, Martin (2010b) "Comparative concepts and descriptive categories in crosslinguistic studies". *Language* 86, 663–687.

Haspelmath, Martin (2016) "The challenge of making language description and comparison mutually beneficial". *Linguistic Typology* 20:2, 299–303.

Haspelmath, Martin, Ekkehard König, Wulf Oesterreicher, and Wolfgang Reible, eds. (2001) *Language typology and language universals, an international handbook.* Berlin: Walter de Gruyter.

Himmelmann, Nikolaus P. (1998) "Documentary and descriptive linguistics". *Linguistics* 36, 161–195.

Himmelmann, Nikolaus P. (2016) "What about typology is useful for language documentation?". *Linguistic Typology* 20:3, 473–478.

Hovdhaugen, Even, Fred Karlsson, Carol Henriksen, and Bengt Sigurd (2000) *The history of linguistics in the Nordic Countries*. Helsinki: Societas Scientiarum Fennica.

Joos, Martin (1957) *Readings in linguistics* I. Chicago: University of Chicago Press.

Kel'makov, Valentin and Sara Hännikäinen (1999) *Udmurtin kielioppia ja harjoituksia.* (Apuneuvoja suomalais-ugrilaisten kielten opintoja varten, 14.) Helsinki: Suomalais-Ugrilainen Seura.

Lazard, Gilbert (1992) "Y a-t-il des catégories interlangagières?". In Susanne Anschütz, ed. *Texte, Sätze, Wörter und Moneme: Festschrift für Klaus Heger*, 427–434. Heidelberg: Heidelberger Orientverlag.

Lehmann, Christian (2001) "Language documentation: a program". In Walter Bisang, ed. *Aspects of typology and universals*, 83–97. (Studia Typologica, 1.) Berlin: Akademie Verlag.

Markus, Elena B., and Fedor I. Rožanskij (2011) *Sovremennyj vodskij jazyk: teksty i grammatičeskij očerk.* Sankt-Peterburg: Izdatel'stvo Sankt-Peterburgsko Instituta Istorii, RAN.

Matteson, Esther (1967) *Bolivian Indian grammars* I–II. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, 16.) Norman: SIL.

Matthews, George Hubert (1965) *Hidatsa syntax.* (Papers on formal linguistics, 3.) The Hague: Mouton.

Miestamo, Matti (2013) "Kielten vertailu kielitypologisessa tutkimuksessa". In Leena Kolehmainen, Matti Miestamo and Taru Nordlund, eds. *Kielten vertailun metodiikka*, 27–55. (Suomalaisen Kirjallisuuden Seuran toimituksia, 1387.) Helsinki: Suomalaisen Kirjallisuuden Seura.

Miestamo, Matti (2016) *Questionnaire for describing the negation system of a language.* Available online at <http://tulquest.huma-num.fr/fr/node/134>. Accessed on 28.10.2018.

Miestamo, Matti, Anne Tamm, and Beáta Wagner-Nagy, eds. (2015) *Negation in Uralic languages*. (Typological Studies in Language, 108.) Amsterdam: Benjamins.

Mühlhäusler, Peter (1996) *Linguistic ecology: language change and linguistic imperialism in the Pacific Region*. London: Routledge.

Nichols, Johanna (2017) *Rich grammar project*. Available online at <https://www.academia.edu/32073015/Rich_Grammar_project>. Accessed on 28.10.2018.

Nickel, Klaus Peter, and Pekka Sammallahti (2011) *Nordsamisk grammatikk*. Kárášjohka: Davvi Girji.

Nickel, Klaus Peter (1994) *Samisk grammatikk*. Karasjok: Davvi Girji.

Nikolaeva, Irina (1995) *Obdorskij dialekt chantyjskogo jazyka*. (Mitteilungen der Societas Uralo-Altaica, 15.) Hamburg: Buske.

Nikolaeva, Irina (2014) *A grammar of Tundra Nenets*. (Mouton Grammar Library, 65.) Berlin: Mouton de Gruyter.

Payne, Thomas E. (1997) *Describing morphosyntax: a guide for field linguists*. Cambridge: Cambridge University Press.

Pope, Kirsten and Máret Sárá (2004) *Eatnigiella: giellaoahpu váldogirji*. Kárášjohka: Davvi Girji.

Sandman, Erika (2013) "Vertailu työvälineenä vähän tutkitun kielen kuvauksessa". In Leena Kolehmainen, Matti Miestamo, and Taru Nordlund, eds. *Kielten vertailun metodiikka*, 56–95. (Suomalaisen Kirjallisuuden Seuran toimituksia, 1387.) Helsinki: Suomalaisen Kirjallisuuden Seura.

Shopen, Timothy, ed. (2007) *Language typology and syntactic description*. 3. Vols. Cambridge: Cambridge University Press.

Siegl, Florian (2013) *Materials on Forest Enets, an indigenous language of Northern Siberia*. (Mémoires de la Société Finno-Ougrienne, 267.) Helsinki: Société Finno-Ougrienne.

Song, Jae Jung, ed. (2011) *Oxford handbook of linguistic typology*. Oxford: Oxford University Press.

Stassen, Leon (1985) *Comparison and universal grammar*. Oxford: Blackwell.

Suihkonen, Pirkko (1995) *Udmurt – English – Finnish dictionary with a basic grammar of Udmurt*. (Lexica Societatis Fenno-Ugricae, 24.) Helsinki: Suomalais-Ugrilainen Seura.

Vasil'ev, Valerian M., and Zinovij V. Učaev (2003) *Marijsko-russkij slovar'*. Joškar-Ola: Marijskoe Knižnoe Izdat.

Wilbur, Joshua Karl (2013) *A corpus-based grammar of spoken Pite Saami*. Doctoral dissertation, Christian-Albrechts-Universität zu Kiel.

Wilbur, Joshua Karl (2014) *A grammar of Pite Saami*. Berlin: Language Science Press.

Winkler, Eberhard (1994) *Salis-Livische Sprachmaterialien*. München: Finno-Ugrisches Seminar.

Winkler, Eberhard (1997) *Krewinisch: zur Erschliessung einer ausgestorbenen ostseefinnischen Sprachform*. (Veröffentlichungen der Societas Uralo-Altaica, 49.) Wiesbaden: Harrassowitz.

Zúñiga, Fernando (2012) "Language description and linguistic typology". In A. Ender, A. Leemann and B. Wälchli, eds. *Methods in contemporary linguistics*, 171–194. Berlin: Mouton de Gruyter.

Аннотация. Матти Миестамо: О взаимоотношении типологической и описательной лингвистики в контексте уральских языков. Лингвисты, занимающиеся типологией, зависят от данных, предоставляемых лингвистами, работающими с конкретными языками. Последние, в свою очередь, используют результаты работы типологов, что позволяет им по-новому взглянуть на материал соответствующего языка. В статье рассматриваются взаимоотношения между типологией и описательной лингвистикой в контексте уральских языков. Положение дел в описании уральских языков оценивается на основе наличия материалов по этим языкам в базе данных Glottolog, широко используемой типологами. В статье обсуждаются способы сближения типологического и описательного подходов в изучении уральских языков, что способствует развитию обоих направлений. В статье даются рекомендации по написанию типологически ориентированных грамматик (уральских) языков. Кроме того, кратко обсуждается использование типологических анкет для описания языков.

Ключевые слова: дескриптивная грамматика, описание языков, документация языков, лингвистическая типология, анкеты, уральские языки

Kokkuvõte. Matti Miestamo: Tüpoloogia ja Uurali keelte kirjeldamise vahelisest suhtest. Keeletüpoloogid sõltuvad andmetest, mis pärinevad üksikkeeltega tegelevatelt deskriptiivses raamistikus töötavatelt keeleteadlastelt. Viimased aga saavad omakorda kasu tüpoloogide töö tulemustest, mis avavad uusi vaatenurki vastavate keelte omaduste kohta. Siinne artikkel käsitleb tüpoloogia ja deskriptiivse Uurali keeleteaduse vahelist suhet. Uurali keelte kirjeldamise seisu uuritakse analüüsides Uurali keeli puudutavate deskriptiivsete allikate kättesaadavust Glottologi andmebaasis, mida tüpoloogid laialdaselt kasutavad. Arutletakse viiside üle, mis aitaksid vastastikkuse kasu eesmärgil lähendada Uurali keelte uurimist ja keeletüpoloogiat, näiteks antakse mitmeid soovitusi tüpoloogilise suunitlusega (Uurali keelte) grammatikate kirjutamiseks. Lõpuks käsitletakse ka tüpoloogilise suunitlusega küsitluskavade kasutamist keelte kirjeldamisel.

Märksõnad: deskriptiivne grammatika, keelekirjeldus, keelte dokumenteerimine, keeletüpoloogia, kirjastamine, küsitluskavad, Uurali keeled