The Uralic languages

PIRKKO SUIHKONEN

This paper deals with the Uralic languages, their regional distribution and relationship with one another. The Uralic languages are spoken in a large area in North and Central Eurasia. Most of the Uralic languages are seriously endangered minority languages – only Finnish, Hungarian, and Estonian are principal national languages spoken in independent countries. Despite being relatives, the Uralic languages differ remarkably from one another. In the west, the Uralic languages have had most intensive relationships with Indo-European languages, and in the east, with Turkic languages. The differences within the language group carry information regarding these contacts. In the research of the Uralic languages, the proto-languages and the original home of the peoples speaking these languages have attracted particular interest. Comparative and historical methods and archaeology have been important in the research of the history of the Uralic languages.

Pirkko Suihkonen, Department of General Linguistics, P. O. Box 9, FIN-00014 University of Helsinki, Finland. E-mail: suihkone@ling.helsinki.fi

Introduction

The Uralic languages, their regional distribution, and relationship with one another are examined in this article. In the first section, the Uralic languages and peoples belonging to the ethnic group identified by the name of the language are presented within the framework of statistical information. The section also contains a short overview of the investigations on the origin and the history of the Uralic languages and their contemporary situation. In the second section of the article, the Uralic languages are described within the framework of language typology. In the last section, some focal points in the study of the Uralic languages will be summarized.

The regional distribution of the Uralic languages

The concept *Uralic languages* is defined on the basis of their genetic classification, which is carried out according to the historical and comparative methods used in historical linguistics. These methods are used to reconstruct a common proto-language (the parent of the contemporary languages) based on the information available from the daughter languages. The so-called *Proto-Uralic* dates back at least 5,000–4,000 years before contemporary era (B.C.E.) (cf. Koivulehto 1999a).

Excluding Hungarian, Estonian, Finnish, Livonian, and most of the Saami languages, the main areas where the Uralic languages are spoken are located in Russia (Fig. 1). The North-Samoyedic languages are spoken in Eurasia's northernmost areas: the Nenets live in the Nenets, Yamalo Nenets, and Taymyr Autonomous Areas, the Enets in the Taymyr Autonomous Area, and the Nganasans in the Taymyr Autonomous Area and the Krasnoyarsk Territory. The Nenets can also be found in the Khanty-Mansiysk Autonomous Area. The Selkup dialects are spoken in the Khanty-Mansiysk Autonomous Area and in the Tomsk Region. In addition to their national areas, speakers of Mari, Udmurt, Komi, and especially Mordvin are scattered across Russia. As a consequence of the relocation of peoples during the Soviet period in Russia, Estonian is also spoken, e.g., in the Krasnoyarsk Territory in Central Siberia. The Karelian Republic, the Tver, Murmansk and Leningrad Regions, and St. Petersburg are the principal areas where Karelians live. The Veps live in the Karelian Republic and in the Vologda and Leningrad

	Size of ethnic group	Estimated number of speakers			Size of ethnic group	Estimated number of speakers
The Baltic-Finnic languages.				12. The Mari languages.	670,868	
1. Finnish	ca. 5,518,115	5,000,000		East Mari	ca. 90 %	
Finland (1999)	ca. 5,100,000	4,788,497		West Mari	ca. 10 %	
Sweden (1992)	11,407					
Estonia (1989)	16,622	5,155		The Permic languages.		
Russia (1989)	67,359	23,274		13. Udmurt	746,793	520,101
2. Estonian				14. Komi Zyrian	344,519	242,515
Estonia (in 1989)	963,281	953,032		15. Komi Permyak	152,060	106,531
Russia (in 1989)	60,363	27,001				
Finland (1999)		10,024		16. The Saami languages.		
3. Karelian (1989)	130,989	62,542		South Saami		300–500
Dvina Karelian	ca.53%			Lule Saami		2,000–3,000
Livvi	ca. 40%			Pite Saami		10
Lude	ca. 7%			Ume Saami		10
4. Veps	12,501	6,355		North Saami		17,000
5. Ingrian	ca. 820	300		Inari Saami		300–500
6. Livonian		< 20		Skolt Saami		300–500
7. Vod		< 20		Akkala Saami		10
				Kildin Saami		650
				Ter Saami		10
The Ugric languages.				Norway	ca.30,000–40,000	20,000
8. Hungarian	ca, 15,000,000	15,000,000		Sweden	ca. 17,000	10,000
Hungary (in 1990)	10,374,823	10,222,529		Finland (in 1999)	ca. 7,000	1,690
Romania (in 1992)	ca. 1,639,135			Russia	ca. 1,900	797 (in 1989)
Slovak Republic (in 1993)	ca. 569,000					
Czech Republic (in 1991)	19,932			The Samoyedic languages.		
Austria (in 1991)	10,556			North Samoyedic languages		
Ukraine (in 1989)	163,111	156,011		17. Nenets	34,665	26,730
Germany (in 1993)	24,240			19. Nganasan	1,278	1,063
The former Yugoslavia	427,00			20. Enets	209	95
The Ob-Ugric languages.				South Samoyedic languages		
9. Khanty	22,521	13,615		18. Selkup	3,612	1,721
10. Mansi	8,474	3,140				
The Finno-Volgaic languages.						
11. The Mordvin languages	1,153,987	773,827				
Erzya	ca. 67%					
Moksha	ca. 33%					

Fig. 1. Peoples and groups of people belonging to the Uralic language family (Rikkinen et al. 1999: 49; Nacionalnyj… 1990; Census of Finland 1999; information on the speakers of the Saami languages is given by Ellen Näkkäläjärvi. Note that linguistic or ethnic data are usually not included in the census information in western European countries. Information on the Hungarian people in the Czech Republic was given by Michaela Kholova (Information Services Unit, CSO), and in Austria, by Ruth Hügelsberger, Statistics Austria, Bundesanstalt Statistik, Österreich).

Regions, and the Ingrians and the Vods are found in the Leningrad Region. The few remaining Livonian-speakers live in northern Latvia. Most of the people belonging to Khanty and Mansi ethnic groups live in the Khanty-Mansiysk Autonomous Area. Khantys also live in the Tomsk Region and the Yamalo-Nenets Autonomous Area.

The indigenous Saami people live in four countries: in Norway, Sweden, Finland, and Russia (Fig. 1). Many speakers of Finnish, Hungarian, and Estonian also live in other European countries as well as in the United States, Canada, and Australia. This migration of the peoples speaking the Uralic languages, particularly those languages with a large number of speakers, is a consequence of numerous factors. The economic factors have been among the most crucial ones. Extensive migration from Finland to the USA, Canada and Australia at the turn and the first half of the twentieth century, and to Sweden in the 1960s and 1970s, are examples of migration caused, to a great extent, by periods of structural transformations and even depression in Finland's economy. Employment opportunities in the fishing industry in northern Norway have been important particularly for people in northern Finland.

Economic reasons caused migration also in the former Soviet Union: the Trans-Baikal railroad and numerous mines in Siberia were built by people who moved to the east and to the north in order to build a brave new world. In the Soviet period, forced transfers of population concerned particularly national minorities and ethnic groups. Migration of Russians to the areas originally populated by minority groups changed permanently the structure of population in these areas. For example, the Estonian and Finnish settlements in Siberia have their roots in this era.

Wars have always caused migration, and a war often reorganizes the distribution of people and languages. Political, economic, and social reasons are also most crucial for a language's death. For example, Livonian can nowadays be considered to be a dead language, but at the beginning of the last millennium Livonians formed a significant minority group in Latvia. Little by little Livonians assimilated to Latvians. Also the fightings that took place in Courland during World War II were disastrous for Livonians. The Livonian fishermen in Courland on the coast of the Baltic Sea were economically less dependent on the Latvian community and thus survived the longest (e.g., Laakso 1991: 116–118).

The degree of endangerment

The contemporary Uralic languages can be divided into four main groups on the basis of the degree of their endangerment. The status of Hungarian, Finnish, and Estonian, which all are principal national languages in independent countries, is protected by law in these countries. The legal-administrative and cultural activities in these countries support and maintain the position of the languages. The languages have the status of majority languages, and their future depends first and foremost on the speakers of these languages.

The position of the second group is less stable. The languages such as Udmurt, Mordvin, Mari, and Komi, which have a relatively high number of speakers, belong to this group. There are various activities within these groups that support their continuing existence. These languages have an official status in the national administrative regions, and they are taught in school at various levels. They are also taught at the universities, in which they are languages of instruction at the departments where the education of teachers and researchers of these languages is arranged. These languages are used in literature, newspapers, and in productions of modern art and popular tradition, and research of these languages is active. That these languages are minority languages in the administrative regions where they are spoken causes serious problems. The number of people belonging to the ethnic groups is usually noticeably higher than that of the native speakers of these languages. In several ways, the position of Karelian, North Saami, and Nenets falls between this second and the third, weaker group. Majority languages have a strong position in everyday life, and, when excluding North Saami, they are, e.g., the languages used in all higher education. Majority languages also have the most important function in administration, and usually all the innovations are learnt through the majority languages spoken in the area. In particular, assimilation of Karelians, Mordvins, and Komi-Permyaks into the main ethnic groups (that usually is Russian) has been most extensive in the twentieth century (Suihkonen 1987; Lallukka 1995). The number of North Saami speakers is much smaller than that of the Volgaic and Permic languages, but the activities supporting North Saami are even more vital than those in the Mari, Mordvin, Udmurt, and Komi Republics. The fact that the speakers of North Saami live in three countries (Norway, Swe-

den, and Finland) complicates cooperation in economic and cultural life. The same kind of situation applies to the Nenets and the Karelians who live in a large area in North Siberia and West Russia, respectively.

The position of the third group is more difficult. The group consists of Veps, and the Ob-Ugric languages Khanty and Mansi (cf. above the North Saami, the Karelians, and the Nenets). The number of speakers is relatively small, as it is with regard to the number of people belonging to the ethnic group. In spite of the fact that many of the languages have official position in the national administrative regions where they are spoken, and that various political and cultural activities support these languages, the press of the majority cultures and languages is strong. The economic situation is difficult, and the structure of economic life does not support minority nationalities. Some parts of the area are rich with natural resources (e.g., one of the world's richest oil fields is located in the Khanty-Mansiysk Autonomous Area), but integration of the local ways of life with the modern industry has proved to be difficult. The small Saami languages in Finland are located between this third group and the fourth group that consists of the small, seriously endangered Baltic-Finnic, Saami and Samoyed languages, some of which are virtually extinct. Documentation on the languages belonging to this group is one of the most urgent tasks of Uralic linguistics. When a language disappears, a huge amount of the mankind's cultural inheritance vanishes. Documentation of these languages is also important because of the languages themselves, as it is the only way to store information on them.

Relationships within the group

The *family tree theory* is one of the most common ways to describe the relationships between the members of the Uralic language group. According to this theory all the branches in the tree represent daughter languages that have developed from one proto-language over the course of the languages' history. In the study of the history of the Uralic languages before the 1990s, Proto-Uralic was dated to circa 6,000–4,000 B.C.E.; Proto-Finno-Ugric 4,000–3,000 B.C.E.; Proto-Finno-Permic 3,000–2,000 B.C.E.; Proto-Finno-Volgaic (Proto-Finno-Mari) 2,000–1,500 B.C.E.; and Proto-Finno-Saami (Early Proto-Baltic-Finnic), the proto-language of the Baltic-Finnic and Saami languages, circa 1,000 B.C.E. (Korhonen 1981: 27; Rédei 2000). According to the most recent studies, the proto-languages in the Uralic branch have a longer history. For example, the break-up of Proto-Finno-Saami is now dated to circa 2,500 B.C.E. (Sammallahti 1998: 33; cf. Koivulehto 1999b). The structure of the family tree is the outcome of a comparative research method. The further one goes back in time, the more difficult it is to recognize linguistic variety.

An alternative approach is given in the *wave theory*, in which the chronological relationships of languages do not have priority. Instead, distribution of various linguistic properties forms areal isoglosses of these properties that are like waves cutting across each other. The relative strength of the boundaries between the isoglosses depends on how they bunch together (see Anttila 1989: 300–309). The family tree and wave theories are complementary to one another: they stress different aspects in the relationships of languages.

According to a third influential approach, the *contact theory*, many of the properties of the contemporary languages are a consequence of contacts between languages over the course of their history (Wiik 2000; cf. Itkonen 1966: 9; Anttila 2000). Evidence of the contacts between the Uralic languages and other languages spoken in the same area in various periods of time can be found in all the Uralic languages. In the west, the Uralic languages have had most intensive contacts with the Indo-European languages and, in the east, with Turkic languages. *Linguistic palaentology*, a research method in which archaeology and linguistics intersect, has had an important role in dating these connections. Linguistic palaentology compares the distribution of cultural words (e.g., the names of plants, seeds, and tools) with the results of archaeological studies and relates the history of the distribution of cultures to linguistic evidence. Linguistic palaentology has also been important in the study of the relationships between the Uralic languages. The earliest connections between the Indo-European and Uralic proto-languages are dated to at least 4,000 B.C.E. According to the theoretical framework of this research, Proto-Uralic was spoken in the area between the Baltic Sea and the middle course of the Volga River, and Proto-Indo-European in the area between the Dnepr and Volga Rivers to the north of the Caspian Sea and the Black Sea. On the basis of the distribution of Indo-Aryan and Indo-Iranian loanwords, the area where Proto-Ugric (the

proto-language of the Ugric languages) and Proto-Samoyed (the proto-language of the Samoyedic languages) were spoken is presumed to have been located to the east of the Ural Mountains (Koivulehto 1999a; Parpola 1999; Kallio 1999).

Indo-European loanwords, particularly those from Pre-Baltic and Pre-Germanic language forms dated to 3,300–2,300 B.C.E., are also important in dating the break-up of Proto-Finno-Saami (Koivulehto 1999b; Sammallahti 1998). In addition to the area in the forest belt between the Baltic Sea and the middle course of the Volga River, the original homeland of the Uralic peoples is thought to have been located in several places between the Altay-Sayan Mountains and northern and central Europe. The most important evidence regarding the areas where the Uralic languages have been spoken is based on archaeological data. Evaluation of archaeological evidence with regard to languages spoken in North and Central Eurasia is still in progress[1].

During the last twenty years, multidisciplinary international study regarding the roots of the peoples speaking the Uralic languages has been productive. In the most recent studies, genetic anthropology has become important. Some of the Uralic-speaking peoples have been studied with regards to their maternal and paternal lineage, that is, the properties of mitochondrial DNA (maternal) and the Y chromosome (paternal). The maternal inheritance of Uralic people living in Scandinavia can be seen as a European subset, and the Saami inheritance as a sub-branch of that, whereas the paternal lineage is linked to some Siberian populations. Among the Yakuts, this Y-chromosome type is very extensive, as it is among the Finns, the Komi, and the Lithuanians. So far, the results of genetic study cover only some fragments of the genetic inheritance of people of North and Central Eurasia (see Sajantila & Pääbo 1995; Savontaus & Lahermo 1999). Interpretation of the results of these studies, as well as those obtained in archaeology or genetic anthropology, is one of the most challenging tasks in the study of the histories of the Uralic languages and peoples.

Typological properties of the Uralic languages

Typologically, languages are defined on the basis of linguistic properties and the relationships between these properties. In the following, some of the structural typological properties used in typological classification of contemporary Uralic languages are presented.

Phonology

Phonologically, there are scarcely any properties common to all the Uralic languages, although there are some common tendencies. One of these is the avoidance of initial consonant clusters: for example, Latin *schola* (school) has become *koulu* in Finnish and *iskola* in Hungarian. In the contemporary languages, this rule no longer obtains. For example, in the southwestern dialect of Finnish (see Raento & Husso 2002: CD-Fig. 1), the initial clusters are part of the phonotactic system of the language. *Vowel harmony* – "a rule whereby all the vowels of a given word must belong to one of a number of partitions of the overall vowel system" (Comrie 1988: 454–455) – is a property that is thought to be inherited from Proto-Uralic. Some variety of palatal vowel harmony is found in the Baltic-Finnic languages, except for Estonian and Livonian. This is true also of West Mari, Hungarian, and some Khanty dialects. The following examples are taken from Finnish and Khanty: in Finnish, *metsä+ssä* signifies 'in a/the forest' (forest+Inessive), *talo+ssa* 'in a/the house' (house+Inessive), and in Khanty's Vakh dialect, *sem+äm* 'my eye' (eye+Possessive-suffix-Singular1), *olv+am* 'my threshold' (threshold+Possessive-suffix-Singular1) (Comrie 1988: 455). Also the South Samoyedic language Kamas, which became extinct in the twentieth century, had partial vowel harmony (Künnap 1999: 11).

Quantity correlation of vowels and consonants is another property considered to be typical of the Uralic languages. The quantity correlation of vowels that has been considered to have its origin in Proto-Uralic, is still represented in Finnish, for example: *tuli* 'fire' : *tuuli* 'wind'. In several languages, the old quantity correlation has been a basis for the development of a new quantity correlation (Itkonen 1966: 61–69; on the history of the development of vowels in the Uralic languages, Sammallahti 1988). *Consonant gradation*, i.e., quantitative and/or qualitative gradation of consonants in the inflected word forms, is considered to be a consequence of parallel developments in the Uralic languages. Consonant gradation is found in the Baltic-Finnic languages (except for Veps and Livonian) and in the Saami languages (except for South Saami). Representative examples

in Finnish are *kukka* : *kuka+n* 'a flower' : 'of a flower' (flower+Genitive) and *pata* : *pada+n* 'a pot' : 'of a pot' (pot+Genitive). In the Baltic-Finnic languages, the gradation concerns stops, whereas in the Saami languages it concerns almost all the consonants (Korhonen 1981: 135–153). In the Samoyedic languages, consonant gradation is found, e.g., in Nganasan (Helimski 1998: 482).

Word initial stress that is thought to have been in Proto-Finno-Ugric, occurs in several Finno-Ugric languages, also in Finnish, but this rule has numerous exceptions in most of the languages. For instance, in Udmurt, the main stress usually falls on the last final syllable, but there are exceptions to this rule for lexical or grammatical reasons. In Erzya, the main stress in the word varies, but it tends to fall on the first syllable (Mosin & Bajuškin 1983: 2). In literary East Mari, the stress is, in most cases, located on the last strong full vowel. In West Mari, the rules of stress placement are more complex: in principle, West Mari has penultimate stress, but there are several exceptions to the rule (Alhoniemi 1985: 17–18). Also some Khanty dialects lack word-initial stress (Kálmán 1976: 1934; cf. Itkonen 1966: 150–159).

Suffixation is the most common type of morphological technique in *word formation* in the Uralic languages. Prefixation is a common morphological method only in Hungarian. The Uralic languages are regarded as *synthetic*, i.e., word forms consist of several elements connected with each other[2]. Synthetic languages express grammatical relationships with affixes, and for this reason word forms can be long. Variation in the synthesis index, developed within the framework of quantitative typology, is remarkable in the Uralic languages. According to the calculations made on the basis of the corpora consisting of one hundred words collected from South Saami, North Saami, Inari Saami, Kildin Saami, Finnish, Erzya, Mari, Udmurt, and Hungarian, the synthesis index (the number of morphemes per number of words) lay between 1.81 (Erzya) and 2.22 (Finnish) (Korhonen 1969: 221). The indices are average figures of the elements they are calculated from. They are based on the structural information on the word form, and they tell nothing about the categories the elements represent.

The Uralic languages are also considered to be *agglutinative*. This means that a grammatical category is always expressed with the same morpheme, and that the morphological elements are conn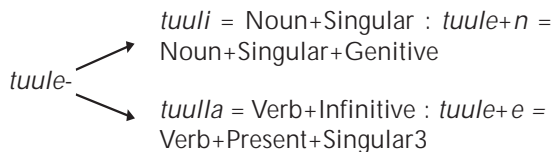ected to the roots one after another, and the boundaries between the elements are clear (see Comrie 1988: 460). The claim that the Uralic languages are agglutinative is only partly true, however: for example in Estonian, Finnish, Hungarian and the Saami languages, the degree of morphophonological variation is noticeable. The range of the indices of agglutination (the number of agglutinative constructions per the number of junctures in the words) calculated for the above-mentioned languages varies between 0.01 (the Saami languages except Inari Saami) and 0.37 (Hungarian) (Korhonen 1969: 221). The opposite technique to agglutination is called *fusion*. In fusional languages, the boundaries between grammatical elements in a word form cannot be clearly drawn. The morphological technique of Skolt Saami is markedly more fusional than most other Uralic languages (Korhonen 1969).

In the nominal inflection of the Uralic languages, there are some elements that have a common origin, but most of the typologically characteristic properties have developed during the separate developments of individual Uralic languages. The number of cases varies between three (Khanty, northern dialects) and twenty-five (Hungarian). In Finnish the number of cases is fifteen, and in North Saami six. In Proto-Uralic, the number of cases is reconstructed to have been at least six (Itkonen 1966: 69). In the basic form, the core case system of the Uralic languages contains grammatical cases (nominative, accusative, and genitive) as well as cases expressing information on locational relations. The nominative form is usually the stem of the word form. The local case systems usually differentiate between cases for static position, motion towards something, and motion away from something. The Baltic-Finnic, Volgaic and Permic languages, and Hungarian further distinguish internal and external local cases. This local case system is illustrated with the following examples from Finnish: *pöydä+ssä* 'in a/the table' (table+Inessive); *pöydä+stä* 'from a/the table' (table+Elative); *pöytä+än* 'to a/the table' (table+Illative); *pöydä+llä* 'on a/the table' (table+Adessive; *pöydä+ltä* 'from a/the table' (table+Ablative); *pöydä+lle* 'on(to) a/the table' (table+Allative) (e.g., Itkonen 1966: 69–78; Korhonen 1992). In the external local cases, Hungarian makes further an additional three-fold distinction between cases that express location close to, and connected with, the element the word form denotes ('to', 'at' and '(away) from' and 'on', 'on(to)' and 'off').

Except for Estonian, all the Uralic languages use suffixes to express *possession*. In the Baltic-Finnic, Volgaic and Saami languages, the case ending precedes the possessive suffix, as in Finnish: *talo+ssa+ni* 'in my house' (house+Inessive+Possessive-suffix-Singular1). In the Ugric languages, the possessive suffix is located before the case suffix, as in Mansi: *kol+uw+n* 'in our house' (house+Lative+Possessive-suffix-Singular1) (Kálmán 1976: 44–46) and in Hungarian: *ház+am+ban* 'in my house' (house+Possessive-suffix-Singular1+Inessive). In the Permic languages and Mari, both of these orders are found. The following examples are from Udmurt: *val+ez+tek* means 'without his/her horse' (horse+Possessive-suffix-Singular3+Caritive), and *gurt+jos+a+z* 'in/to his/her villages' (village+Plural+Inessive/Illative+Possessive-suffix-Singular3) (Comrie 1988: 464). All the Uralic languages have postpositions, and, in some languages, they can be used as prepositions, as in Finnish: *pöydä+n keske+llä* : *keske+llä pöytä+ä* 'in the centre of the table' (table+Genitive centre+Allative : centre+Allative table+Partitive).

In expressing *number*, all the Uralic languages distinguish singular and plural, and the Saami, Ob-Ugric and Samoyedic languages also a dual. In some languages, the number distinction of possessive suffixes concerns both possessors and the possessed. Particularly the third person possessive suffixes have developed to express other grammatical categories. For example in Udmurt, the third person possessive suffixes have developed nominalizers that are used in defining a delimited, closed set: *mehaṅik+jos+my+ly+os+yz* 'those which belong to our mechanics' (mechanic+Plural+Possessive-suffix-Plural1+Dative+Plural+Nominalizer) (Suihkonen 1990: 297–301, 305–307; cf. Kálmán 1976: 69–70 on Mansi; Nikolaeva 1999: 80–84 on Khanty). Hungarian is the only Uralic language in which the distinction between *indefinite* and *definite noun phrases* is systematically done with the aid of articles: *egy ház* 'a house', *a ház* 'the house', *ez a ház* 'this (the) house'. In Mordvin, the definiteness of noun phrases is expressed with a separate definite declension: *kudo* 'a house', *kudo+ś* 'the/that house' (Mosin & Bajuškin 1983: 24). It must also be noted that grammatical *gender* does not exist in the Uralic languages, although nominal derivation does admit distinctions of natural gender. For example in Finnish, *kirjailija,* 'a writer (male or female)', and *kirjailija+tar,* 'a female writer', can be distinguished.

The borders between the *parts of speech* are not always clear in the Uralic languages. In all of them, there are many words that can be inflected both nominally and verbally. A Finnish example is the word stem *tuule-* 'wind', 'to blow' (Itkonen 1966: 80):

tuule- → *tuuli* = Noun+Singular : *tuule+n* = Noun+Singular+Genitive

tuule- → *tuulla* = Verb+Infinitive : *tuule+e* = Verb+Present+Singular3

These words belong to the ancient vocabulary of the Uralic languages. The distinction between nouns and adjectives is weak, as is that between adjectives and adverbs. The distinction between nouns and adjectives is differentiated furthest in the Saami languages, in which most of the adjectives have different forms in the prenominal and the predicative position. In addition to adjectives, also some nouns can take the comparison suffix as in Finnish: *ranta* : *ranne+mpa+na* 'shore' : 'closer to the shore' (shore+Comparative+Essive). An interesting property typical of all these languages is a large number of nominal and adverbial forms of verbs.

All the Uralic languages distinguish *person* and *number* in *verbal inflection*. The personal endings are connected to the verbs after the modal and temporal suffixes. In the compound tense forms the personal suffixes are often connected with auxiliaries that can also contain information on temporal and modal distinctions. In number, the Saami, Ugric and Samoyedic languages distinguish a dual in addition to a singular and a plural. The Uralic languages differ with regard to the types of *tense* and *mood* expressed with verbal inflection. The basic distinction in expressing tense concerns the opposition between present and past, or past and non-past, but the variety within these categories is remarkable. Grammatical future is found in Komi, Udmurt, Hungarian, and in some Khanty dialects. In Udmurt, the future is expressed synthetically with the suffix *-o-*: *jual+o+d* 'you will ask' (ask+Future+Singular2), *mis'k+o+dy* 'they will wash' (wash+Future+Plural2) (Perevoščikov et al. 1962: 201). In Hungarian, the verb *lesz* (to be) has the future form: *lesz+ek* 'I shall be', etc. From the other verbs, the future is formed with the present tense of the auxiliary *fog* (start) and the infinitive form of the verb that carries the lexical meaning of the verbal

phrase: *vár+ni fog+nak* 'they will wait' (wait+Infinitive will+Plural3). In practice, the compound future is no longer used frequently (Keresztes 1974: 49).

Some tense markers that have been reconstructed to be Proto-Finno-Ugric are represented in the verbal morphology of the modern Finno-Ugric languages. Also some elements of expressing mood (imperative and conjunctive/subjunctive suffixes) from the same historical period have their representatives in the contemporary languages. All the Uralic languages have the indicative and the imperative moods, and all the Finno-Ugric languages (except for Mansi) have the conditional. Various types of evidentiality are expressed grammatically in several Finno-Ugric languages. In Estonian, there is a special mood used in these kinds of expressions (indirect mood, *Modus obliquus*): *sa luge+vat* signifies 'it is said that you read' (you read+Indirect-mood+Present) (Laanest 1982: 239, 265). In Udmurt, some tense forms also carry information on evidentiality: *bödti+ško völ+em* 'it was said that I was finishing' (finish+Present+Singular1 Auxiliary+Perfect) (Fokos-Fuchs 1954: 154). The indicative, the basic mood, is not marked in the Uralic languages.

In the Baltic-Finnic languages, the aspectual distinctions are involved with case marking, e.g., when the activity is not terminative, the noun is in the partitive form, as in Finnish: *rakenna+n talo+a* 'I am building a house' (build+Singular1 house+Partitive). When the activity is terminative, the noun is in the genitive form: *rakenna+n talo+n* 'I build a/the house' (build+Singular1 house+Genitive). In the case ending *-n* above, which is analysed as genitive in the modern Finnish, the old Proto-Uralic genitive case **-n* and the accusative case **-m* have fallen together. In Hungarian, imperfective verbs are changed to perfective ones with perfective prefixes: *a barát+om+hoz men+t+em* 'I was going to my friend' (Definite-Article friend+Possessive-Suffix+Singular1+Allative go+Past+Singular1); *el+men+t+em a barát+om+hoz* 'I went to my friend' (Perfective-prefix+go+Past+Singular1 Definite-article friend+Possessive-suffix-Singular1+Allative) (Kerestesz 1974: 114–115). The progressive aspect is typically expressed with verbal nominal forms inflected with locative cases: in Finnish, *tyttö o+n kävele+mä+ssä* 'a girl is walking' (girl be+Singular3 walk+Infinitive3+Inessive). Also verbal derivation is used in expressing aspectual distinctions as in Udmurt: *šonty+ny* 'to wave' (terminative): *šona+ny* 'to wave' (non-terminative; *-ny* = Infinitive). Information on aspect is given, e.g., with the definite verbal conjugation in the Ugric and Samoyedic languages and Mordvin (Hajdú 1988: 16–17; Mosin & Bajuškin 1983: 95–105). In the definite verbal conjugation verbal inflection is used to give information on the person and number of the subject and object (or, the agent and the target). The following examples are from Khanty's Muzhi dialect: *mạ+s+em* stands for 'I gave it' (give+Past+Singular1/Singular3), *mạ+s+lam* 'I gave those two / those' (give+Past+Singular1-Dual3/Plural3) (Rédei 1965: 66–67).

The variety in expressing *diathesis* can be compared with that of expressing aspect [*diathesis*: "the sense is that of the role or 'placing' of a subject, e.g. as agent in relation to an active verb, or as patient or 'undergoer' in relation to a passive." (Matthews 1997: s. v. *diathesis*)]. A personal passive is found in the Ob-Ugric languages: in Khanty's Muzhi dialect, *-aj-* is the passive suffix, and *kit+s+aj+ən* means 'I was sent' (send+Past+Passive+Singular1), *kit+s+aj+mən* 'we two were sent' (send+Past+Passive+Dual1), and *kit+s+aj+uw* 'we were sent' (send+Past+Passive+Plural1) (Rédei 1965: 71). Finnish has the impersonal passive, in which the personal ending of the passive form contains information on the agent that is not specified: *talo rakenne+tt+i+in* 'the house was built' (house build+Passive+Past+Unspecified-agent). The verbal nominal system (e.g., participles) usually contains passive and active forms. In Udmurt, the participle *-emyn* often has also the passive function: *pinal+ez vande+myn…* 'her child was stabbed…' (child+Possessive-suffix+Singular3 stab+Participle+Past+Passive) (Fokos-Fuchs 1954: 160). In the Uralic languages, also some verbal derivational suffixes have passive and reflexive function – cf. the examples from Finnish: *kulke+a* 'to go' (go+Infinitive): *kulke+utu+a* 'be carried (conveyed); drift, be driven (with the wind)' (go+Reflexive-passive-derivation-suffix+Infinitive); *puke+a* 'dress' (dress+Infinitive): *puke+utu+a* 'dress oneself' (dress+Reflexive-passive-derivational-suffix+Infinitive). Reflexivity is one of the categories that are typically expressed with verbal derivation in the Uralic languages: in Mansi, the suffix *-xat-* has the reflexive function: *masi* 'to clothe': *mas+xat+i* 'to clothe oneself' (Kálmán 1976: 56). Also causativity and double causativity in the Uralic languages are expressed by the verbal derivation, as in Finnish: *kulke+a* 'to go' (go+Infinitive): *kulje+tta+a* 'to transport; to carry;

to take' (go+Causative+Infinitive): *kulje+t+utta+a* 'let/make smbd transport, carry, take smthg' (go+Causative+Causative+Infinitive) (*-a* = Infinitive suffix).

With some exceptions, the use of the conjugated negative verb is typical of the Uralic languages, although other kinds of techniques of expressing *negation* exist as well. In Finnish, *minä e+n tule* signifies 'I do not come' (I not+Singular1 come+Connegative-form), s*inä e+t tule* 'you do not come' (you not+Singular2 come+Connegative-form). In Estonian, negation is not expressed with finite negative verb, but information on the person is given with personal pronouns: *ma ei palu* 'I do not ask' (I not ask); *sa ei palu* 'you do not ask' (you not ask). Person and number are distinguished in the prohibitive forms (Laanest 1982: 242–271). Prohibition and sentential negation are given with different lexical elements, e.g., in the Ugric languages, which have different particles in expressing prohibition and negation (Mansi: *taw at kāsalaste* 's/he did not note it', *tuw ul minen!* 'do not go there!') (Kálmán 1976: 67–68).

The syntactic level

The Uralic languages also differ from one another *syntactically*. In Proto-Uralic, the *order of the main syntactic constituents*, O(bject) and V(erb, predicate), is reconstructed as OV, and in Proto-Finno-Ugric, when taking S(ubject) into account, SOV (Raun 1988: 568–569). In the modern Uralic languages this order varies. The principal word order of neutral statements in Udmurt is SOV, but, for instance, in the Baltic-Finnic languages it is SVO. Thus, in Udmurt, *ta piosmurt val+ez vi+i+z* 'this man killed the horse' (This man horse+Accusative kill+Past+Singular3), but in Finnish, *minä ota+n kirja+n* 'I take the book' (I take+Singular1 book+Genitive). In the Uralic languages, the subject of the sentence is typically in the nominative form, and the object of the transitive sentences is distinguished with the case marking. The agent of the passive sentences in Khanty is in the locative case form. In Finnish, the agent of the passive type sentences, in which the predicate is formed by the III infinitive, is in the genitive form: *taulu o+n Peka+n maalaa+ma* 'the picture is painted by Pekka' (picture+Nominative be+Present+Singular3 Pekka+ Genitive paint+Infinitive-III+Nominative).

In basic *nominal sentences*, several Uralic languages have a copula. Finnish is one of them: *talo o+n suuri* 'the house is big' (house be+Singular3 big). In some languages, in the present tense, this sentence type also occurs without a copular verb, as in Udmurt: *aršin kuź* 'aršin is long' (aršin long; *aršin* is an old measure of length) (Perevoščikov et al. 1962: 137; cf. Comrie 1988: 473). Nouns can be conjugated in nominal and locative sentences in some Samoyedic languages and Mordvin. The following examples are from Erzya: *od* 'young', *od+an* 'I am young' (young+Singular1), *od+tado* 'you are young' (young+Plural2); *tese* 'here', *tes+an* 'I am here' (here+Singular1), *tes+at* 'you are here' (here+Singular2) (Mosin & Bajuškin 1983: 4). The Uralic languages differ from one another with regard to the order of the comparative standard and adjective, and the use of comparative conjunctions. In Finnish, the comparative conjunction *kuin* 'than' is used: *Markus o+n vanhe+mpi kuin Matias* 'Markus is older than Matias' (Markus be+Singular3 old+Comparative-Suffix than+Comparative-Conjunction Matias=Comparative-standard). In Udmurt, instead of the comparative conjunction, the comparative standard is marked by the ablative case ending: *metr aršin+leś kuź+gem* 'the meter is longer than aršin' (meter aršin+Ablative=Comparative-standard long+Comparative-Suffix) (Perevoščikov et al. 1962: 137). Conjunctions in the modern Uralic languages are loans or they have developed after the Proto-Uralic period. In Proto-Uralic, subordination was expressed with non-finite verbal forms (Korhonen 1981: 346). Asyndetic co-ordination is still used, e.g., in Mansi (cf. Kálmán 1976: 346), and coordinative conjunction is not obligatory in some structures, e.g., in Finnish: *toinen auto ol+i sininen*, *toinen* (*ol+i*) *punainen* 'one car was blue, the other one (was) red' (one car be+Past+Singular3 blue, other-one (be+Past+Singular3) red). Also the instructive case connected with nouns has the coordinative function, e.g., in adverbial phrases: *kävele+n aamu+i+n illo+i+n* 'I walk every morning and evening' (walk+Singular1 morning+Plural+Instructive evening+Plural+Instructive).

In the *noun phrases*, the determiner, adjective attribute, and numeral precede the noun, as in Hungarian: *piros alma* 'a red apple' (red apple); *két barát* 'two friends' (two friend+Singular), *ez a ház*: 'this house' (this Definite-Article house) (Keresztes 1974: 145). In the noun phrases consisting of a numeral and a noun, the noun is in the singular form also when the value of the numeral is more than one, as in Udmurt: *vit nunal* 'five days' (five day+Singular+Nominative) (cf. the above example

in Hungarian). In Finnish, the noun is in the partitive case: *kaksi päivä+ä* 'two days' (two day+ Singular+Partitive). In most of the Uralic languages, the determiner and adjective do not agree with the noun, although they do, for instance, in the Baltic-Finnic languages such as Finnish: *tä+ssä piene+ssä kaupungi+ssa* 'in this small town' (this+ Inessive small+Inessive town+Inessive).

In linguistic typology, the order of the noun modifiers and the main syntactic constituents [(S)ubject, V(predicate), (O)bject] has been one of the characteristic parameters in defining language types. Realization of the word order combinations within noun phrases and sentences has been linked with the occurrence of adpositions in languages and used in defining word-order universals (Greenberg 1980; Hawkins 1983: 74, 67, 79, 81). The pre-nominal order is typical of SOV languages, which is the reconstructed order of Proto-Uralic. It is claimed that the surface word order of the Uralic languages is relatively free owing to their high degree of synthesis. Because of the variety in the word order patterns in the Uralic languages it is also claimed that (at least in the Uralic languages spoken in Europe) there is no word order model that can be said to be typical of them all (Vilkuna 1998). Word-order typology as such is not exhaustive in describing the syntactic relationships of languages. In the Uralic languages, grammatical relations are expressed with dependent morphological elements and, quite often, the word order is used to express thematic relations. Typology developed for investigating the organization of syntactic elements on the basis of discourse-pragmatic reasons can be better used in characterizing the free word order languages, such as Hungarian and Finnish (see Sasse 1995).

Summary

The relationships of the Uralic languages are defined with the aid of genetic classification. On the basis of historical linguistic studies and by using typological information on languages we can follow the processes that take place when a language is changing. In spite of the fact that the historical documents of the Uralic languages are relatively young – the oldest documents, from Hungarian and Karelian, date from the thirteenth century – many historical properties of the Uralic languages are known. On the basis of the typological description of the Uralic languages it can be said that the differences between individual Uralic languages are even more notable than expected. An interesting question is how close the contemporary Uralic languages are to one another typologically. The Uralic languages are not isolated: over the course of history they have been influenced by numerous other languages and cultures. "The Uralic languages belong to a large areal-typological linguistic group extending from the Atlantic to the Pacific, also including the Turkic, Mongolic and Tungusic languages. The members of these four language families are characterised by a number of common structural features" (Korhonen 1992: 163). Another interesting and important question is how close the Uralic languages are typologically to the other languages spoken in North and Central Eurasia.

The Uralic languages are spoken in a large area in Europe and North and Central Asia. Archaeological studies have provided us with information on the material cultures of the area. The rapid advances in genetic science have produced a great amount of data on the genetic inheritance of the peoples living in this area. There still remain numerous unanswered questions about the history of the Uralic languages and peoples. The use of multidisciplinary methods in trying to find answers to these questions is not only promising and challenging, but also very complex, and usually the methods used in other fields of science shed no direct light on the history of languages. Also, interactions between the research methods used in comparative historical studies and multi-disciplinary studies are waiting for new evaluations. The more information on the history of the Uralic languages and the other languages spoken in the same area will be available, the better the relationships between these languages can be understood.

NOTES

[1] On different aspects on the history of the Uralic people and Uralic languages, see Itkonen 1966: 22–31. A sum-

mary of the theories concerning the original homeland (*Urheimat*) of the Uralic peoples is given in Rédei (2000). For a complete summary of the most recent studies on the connections of the Indo-European and Baltic-Finnic languages, see Anttila (2000).

[2] Historically, the traditional morphological typological classification of languages, based mostly on Edward Sapir's work in the 1920s (see Sapir 1949), has its roots in the German language typology (*Sprachtypology*) from the nineteenth century. In this classification, *isolation*, *agglutination*, and *fusion* are the main classes of the morphological technique. These techniques are specified with the properties derived from other levels of languages. Moreover, it has been considered that the development of the morphological technique of languages is cyclic.

REFERENCES

Alhoniemi A (1985). Marin kielioppi. *Apuneuvoja suomalais-ugrilaisten kielten opintoja varten* X. Suomalais-Ugrilainen Seura, Helsinki.

Anttila R (1989). Historical and comparative linguistics. *Amsterdam Studies in the Theory and History of Linguistic Science, Series IV.* 1st ed. 1972. John Benjamins, Amsterdam.

Anttila R (2000). The Indo-European and Baltic-Finnic interface: time against the ice. In Renfrew C, A McMahon & L Trask (eds.) *Time depth in historical linguistics*, 421–528. The McDonald Institute for Archaeological Research, Cambridge.

Census of Finland in 1999 (1999). Statistics Finland. Autumn, 2001. <*http://statfin.stat.fi/statweb/statfincatalog_Väestölaskenta.asp*>

Comrie B (1988). General features of the Uralic languages. In Sinor D (ed). The Uralic languages. Description, history and foreign influences. *Handbook of Uralic Studies*, 451–477. EJ Brill, Leiden.

Fokos-Fuchs DR (1954). Grammaticalischer Abriss. In Wichmann Y (ed). Wotjakischer Chrestomathie mit Glossar. *Apuneuvoja suomalais-ugrilaisten kielten opintoja varten* II, 135–166. Suomalais-Ugrilainen Seura, Helsinki.

Greenberg JH (1980). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg JH (ed). *Universals of language*, 73–113. 1st ed. 1963. MIT Press, Cambridge.

Hajdú P (1988). Die samojedischen Sprachen. In Sinor D (ed). The Uralic languages. Description, history and foreign influences. *Handbook of Uralic Studies*, 3–40. EJ Brill, Leiden.

Hawkins JA (1983). *Word order universals*. Academic Press, New York.

Helimski E (1998). Nganasan. In Abondolo D (ed). *The Uralic languages*, 480–515. Routledge, London.

Itkonen E (1966). Suomalais-ugrilaisen kielen- ja historiantutkimuksen alalta. *Suomalaisen Kirjallisuuden Seura, Tietolipas* 20.

Janhunen J (2000). Reconstructing Proto-Uralic typology spanning the millennia of linguistic evolution. In Nurk, Anu, Triinu Palo & Tõnu Seilenthal (eds). *Congressus nonus internationalis Fenno-Ugristarum, 7.–13.8.2000, Tartu.* Pars I. *Orationes plenariae & Orationes publicae*, 59–76. Tartu.

Kallio P (1999). Varhaiset indoeurooppalaiskontaktit. In Fogelberg P (ed). Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan. *Bidrag till kännedom av Finlands natur och folk* 153, 237–239.

Kálmán B (1976). *Chrestomathia vogulica.* Tankönyvkiadó, Budapest.

Keresztes L (1974). *Unkarin kieli.* SKS, Helsinki.

Koivulehto J (1999a). Varhaiset indoeurooppalaiskontaktit: aika ja paikka lainasanojen valossa. In Fogelberg P (ed). Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan. *Bidrag till kännedom av Finlands natur och folk* 153, 207–236.

Koivulehto J (1999b). Verba mutuata. Quae vestigia antiquissimi cum Germanis aliisque Indo-Europaeis contactus in linguis Fennicis reliquerint (Ed. KPh Ruppel). *Suomalais-Ugrilaisen Seuran toimituksia* 237.

Korhonen M (1969). Die Entwicklung der morphologischen Methode im Lappischen. *Finnisch-ugrische Forschungen* XXXVII, 203–362.

Korhonen M (1981). Johdatus lapin kielen historiaan. *Suomalaisen Kirjallisuuden Seuran toimituksia* 370.

Korhonen M (1992). Remarks on the structure and history of the Uralic case system. *Journal de la Société Finno-Ougrienne* 82, 163–180.

Künnap A (1999). Kamass. *Languages of the world, Materials* 185. Lincom Europa, München.

Laakso J (1991). Itämerensuomalaiset sukukielemme ja niiden puhujat. In Laakso J (ed). *Uralilaiset kansat. Tietoa suomen sukukielistä ja niiden puhujista*, 49–122. WSOY, Porvoo.

Laanest A (1982). *Einführung in die ostseefinnischen Sprachen.* 1st ed. 1975. Helmut Buske, Hamburg.

Lallukka S (1995). Assimilation and its measurement: Finno-Ugrian peoples of Russia. In Leskinen, H, P Pulkkinen, R Raittila & T Seilenthal (eds). Conressus octavus internationalis Fenno-Ugristarum, Jyväskylä 10–15.8.1995. Pars I. *Orationes plenariae et conspectus quinquennales*, 83–95. Moderatores, Jyväskylä.

Mosin MV & NS Bajuškin (1983). Ersämordvan oppikirja. *Apuneuvoja suomalais-ugrilaisten kielten opintoja varten* VIII. Suomalais-Ugrilainen Seura, Helsinki.

Nacionalnyj sostav naselenija RSFSR po dannym vsesojuznoj perepisi naselenija 1989 g. Gosudartsvennyj komitet RSFSR po statistike (1990). Respublikanskij informacionno-izdatel'skij centr, Moskva.

Nikolaeva I (1999). Ostyak. *Languages of the world, Materials* 305. Lincom Europa, München.

Parpola A (1999). Varhaisten indoeurooppalaiskontaktien ajoitus ja paikannus kielellisen ja arkeologisen aineiston perusteella. In Fogelberg P (ed). Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan. *Bidrag till kännedom av Finlands natur och folk* 153, 180–206.

Perevoščikov PN, VM Vahrušev, VI Alatyrev, AA Pozdeeva & IV Tarakanov (eds) (1962). *Grammatika sovremennogo udmurtskogo jazyka. Fonetika i morfologija.* Udmurtskoe knižnoe izdatel'stvo, Izhevsk.

Raento P & K Husso (2002). Cultural diversity in Finland. *Fennia* 180, 151–164.

Raun A (1988). Proto-Uralic comparative, historical morphosyntax. In Sinor D (ed). The Uralic languages. Description, history and foreign influences. *Handbook of Uralic Studies*, 555–571. EJ Brill, Leiden.

Rédei K (1965). Northern Ostyak chrestomathy. *Indiana University Publications, Uralic and Altaic Series* 47.

Rédei K (2000). Urheimat und Grundsprache (Wissenschaftliche Hypothesen und unwissenschaftliche Fehlgriffe). In Nurk A, T Palo & T Seilenthal (eds). Congressus nonus internationalis Fenno-Ugristarum, 7.–13.8.2000, Tartu. Pars I. *Orationes plenariae & Orationes publicae,* 109–126. Tartu.

Rikkinen K, P Maaranen & P Suihkonen (1999). Erätaloudesta nyky-yhteiskuntaan. In Westerholm J & P Raento (eds). *Suomen kartasto*, 48–53. Suomen Maantieteellinen Seura & WSOY, Helsinki.

Sajantila A & S Pääbo (1995). Language replacement in Scandinavia. *Nature Genet* 11, 359–360.

Sammallahti P (1988). Historical phonology of the Uralic languages. In Sinor D (ed). The Uralic languages. Description, history and foreign influences. *Handbook of Uralic Studies*, 478–554. EJ Brill, Leiden.

Sammallahti P (1998). Saamic. In Abondolo D (ed.) *The Uralic languages*, 43–95. Routledge, London.

Sapir E (1949). *Language. An introduction to the study of speech.* 1st ed. 1921. Harcourt, Brace & World, New York.

Sasse HJ (1995). Prominance typology. In Jacobs J, W Sternefeldt & T Vennemann (eds). Syntax. *Ein internationales Handbuch zeitgenössischer Forschung/An International Handbook of Contemporary Research* 2, 1065–1075. Walter de Gruyter, Berlin.

Savontaus M-L & P Lahermo (1999). Uralilainen muinaisuutemme väestögenetiikan valossa. In Fogelberg P (ed). Pohjan poluilla. Suomalaisten juuret nykytutkimuksen mukaan. *Bidrag till kännedom av Finlands natur och folk* 153, 60–63.

Suihkonen P (1987). Suomensukuiset kansat väestötilastojen valossa. *Terra* 99, 209–232.

Suihkonen P (1990). Korpustutkimus kielitypologiassa sovellettuna udmurttiin (Computer corpus analysis in language typology applied to Udmurt). *Suomalais-Ugrilaisen Seuran toimituksia* 207.

Viitso T-R (2000). Finnic affinity. In Nurk A, T Palo & T Seilenthal (eds). Congressus nonus internationalis Fenno-Ugristarum, 7.–13.8.2000, Tartu. Pars I. *Orationes plenariae & Orationes publicae*, 153–178. Tartu.

Vilkuna M (1998). Word order in European Uralic. In Siewierska A (ed). Constituent order in the languages of Europe. Empirical approaches to language typology. *Eurotyp* 20: 1, 173–233. Mouton de Grueter, Berlin.

Wiik K (2000). European lingua francas. In Künnap A (ed). The roots of peoples and languages of Northern Eurasia II & III. *Fenno-Ugristica* 23, *Historica Fenno-Ugristica*, 202–236. Tartu.