

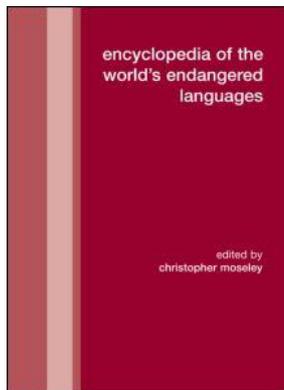
This article was downloaded by: 10.3.98.104

On: 02 Jun 2020

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



## encyclopedia of the world's endangered languages

christopher moseley

### Europe and North Asia

Publication details

<https://www.routledgehandbooks.com/doi/10.4324/9780203645659.ch3>

Tapani Salminen

**Published online on: 01 Mar 2007**

**How to cite :-** Tapani Salminen. 01 Mar 2007, *Europe and North Asia from: encyclopedia of the world's endangered languages* Routledge

Accessed on: 02 Jun 2020

<https://www.routledgehandbooks.com/doi/10.4324/9780203645659.ch3>

**PLEASE SCROLL DOWN FOR DOCUMENT**

Full terms and conditions of use: <https://www.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

---

## Europe and North Asia

*Tapani Salminen*

---

### An outline of the chapter

The coverage of this chapter is defined geographically with certain ethnolinguistic considerations. In broad terms, it includes languages spoken in Europe (including Malta and Cyprus), Anatolia, Caucasia, Central Asia, Mongolia, Manchuria and Siberia. The exact border line is drawn so that the contiguous areas of Turkic and Mongolic languages are included in this chapter while Indo-Iranian and Semitic languages spoken on the border regions belong to adjacent chapters. This means, for instance, that languages spoken in Tajikistan are not represented here insofar as they are Iranian, even though Tajikistan is one of the Central Asian countries. Eastern Turkestan (Sinkiang) and Inner Mongolia are largely included in this chapter while Tibet is not. Enclaves of Indo-Iranian and Semitic languages within the above-mentioned geographical boundaries are, however, dealt with here, and Turkic and Mongolic languages spoken inside the areas of other chapters are found there. All Eskimo-Aleut languages, including those spoken in Siberia, are dealt with in Chapter 1 on North America, and the entry on Ainu is found in Chapter 5 on East and Southeast Asia. The special cases as well as the overall contents of the chapter are explained in the following list of language families, which also serves as an index to the subsequent survey of linguistic history and diversity of the area:

- 1 the *Basque* language;
- 2 the *Uralic* (or Finno-Ugrian) language family;
- 3 the *Indo-European* language family *except* the Indo-Iranian branch (see Chapter 4); the Romani and Lomavren languages of the Indic subbranch as well as the Ossete language and the Tat languages of the Iranian subbranch are, however, included in this chapter;
- 4 the Cypriot Arabic and Maltese languages of the Arabic branch of the *Semitic* language family (see Chapters 4 and 7 for the other Semitic languages);
- 5 the *Kartvelian* language family;
- 6 the *Abkhaz-Adyge* language family;

- 7 the *Nakh-Daghestanian* language family;
- 8 the *Turkic* language family *except* the Khalaj and Kashkay languages (see Chapter 4) and the Salar and Saryg Yugur languages (see Chapter 5);
- 9 the *Mongolic* language family *except* the Mogol language (see Chapter 4) and the Shira Yugur language and the Monguor branch (see Chapter 5);
- 10 the *Tungusic* language family *except* the Manchu branch (see Chapter 5);
- 11 the *Nivkh* language;
- 12 the *Yeniseian* language family;
- 13 the *Yukagir* language family;
- 14 the *Chukotkan* language family;
- 15 the *Kamchatkan* language family.

Languages belong to the same language family, if the similarities among them, as formulated by Watkins in *The Indo-European languages* (1998: 26), ‘require us to assume they are the continuation of a single prehistoric common language’. Watkins is discussing Indo-European, but obviously, any group of languages must meet this definition to be called a language family. The above classification can be regarded as very firm in this respect, none of the families being in any way controversial. Notably, the genetic connection between the Nakh languages and the Daghestanian languages, that is, their belonging to a single Nakh-Daghestanian language family, has been definitively demonstrated (see Nichols 2003).

On the other hand, no solid evidence has been presented for the genetic affinity between Abkhaz-Adyge and Nakh-Daghestanian, although such a ‘North Caucasian’ family is propagated by people involved in long-range comparisons and frequently quoted in general handbooks. Another alleged macro-family, typically included in overall classifications, often even without a question mark, is, of course, ‘Altaic’, consisting of Turkic, Mongolic and Tungusic, and extending to Korean and Japanese in more extreme versions. Furthermore, a ‘Chukotko-Kamchatkan’ family, incorporating Chukotkan and Kamchatkan, is often taken for granted, and other combinations such as Uralic plus Yukagir are occasionally called language families as well. In all of these cases, it is possible and indeed plausible to explain the shared features on the basis of areal influence and typological similarity rather than genetic affinity, and the extant material rather supports the idea of ancient contacts between genetically unrelated languages. The basic issue here is that these groups contrast markedly with established language families such as Uralic and Indo-European which, although internally diverse, can only be explained as deriving from a single proto-language. For a recent attempt to establish a genetic connection between Chukotkan and Kamchatkan, see Fortescue 2005.

In the particular case of Altaic, the best recent elaboration of the plausibility of the areal rather than genetic explanation actually comes from a firm supporter of the genetic hypothesis. Dybo in the entry on the ‘*altajskie jazyki*’ in the *Jazyki Rossijskoj Federacii i sosednih gosudarstv I* (1997: 79–86) presents a large number of lexical comparisons which according to her represent cognate words inherited from Proto-Altaic. Two things are, however, conspicuously missing. First, while there are several words related to fundamental concepts such as body parts in the Altaic corpus, the most basic words in the relevant semantic fields remain restricted to a single family without cognates in the other Altaic families. Second, the material in the Altaic corpus shows only shallow phonological differentiation of shared lexicon, which cannot be the case of words deriving from Proto-Altaic, as the alleged proto-language, if valid,

must have gone through a drastic series of changes not only in morphology and lexicon but also in its sound system. Schönig in the chapter on ‘Turko-Mongolic relations’ in *The Mongolic languages* (2003) discusses the matter further.

While ‘North Caucasian’, ‘Altaic’ and ‘Chukotko-Kamchatkan’ appear in literature both as presumed genetic and factual areal units, there are other labels that only have a conventional meaning. For instance, it is a matter of consensus that the term ‘Caucasian languages’ only refers to a geographic group including the Kartvelian (South Caucasian), Abkhaz-Adyge (Northwest Caucasian) and Nakh-Daghestanian (Northeast Caucasian) languages, as apparently no one would suggest that the Kartvelian family was related to the others. Similarly, ‘Paleoasiatic (or Paleosiberian) languages’ is only used as a cover term for Chukotkan, Kamchatkan, Nivkh, Yeniseian and Yukagir, because they simply represent languages not belonging to any larger family in Siberia.

The one and only language in the area that can be called genuinely mixed, Copper Island Aleut, based on Aleut and Russian, as explained by Golovko in the *Mixed languages* (1994: 113–21), in its mixed nature second only to Michif, described by Bakker in the same volume (13–33), is dealt with in Chapter 1.

There are several other languages that can be called mixed with regard to their lexicon, although their grammar derives from a single language. Examples are Town Frisian, which can be described as a variety of Dutch heavily influenced by West Frisian, and Äynu in Sinkiang, which is an in-group language whose speakers use it besides their native Uygur (see Wurm 1997). A whole group of languages are those with Romani lexicon incorporated into the grammar of local languages, including, in Europe, Angloromani, Scandinavian Traveller languages, Caló or Hispanoromani, Dortika or Hellenoromani, and Šatrovački in the central Balkans. The earlier English, Scandinavian, Iberian and Greek dialects of Romani appear to be extinct but their lexical resources serve to form these special varieties. Boshá, in Caucasia, is similarly based on Armenian, with influences from Lomavren, an extinct language closely related to Romani. Furthermore, there are a number of secret or in-group languages such as Hantyrka, Rotwelsch and Yeniche in Central Europe, Polari and Shelta (Cant) or Yola in the British Isles, and Quinqui in Spain. A few pidgin languages such as Taymyr Pidgin (Govorka) that is still spoken on Taymyr peninsula, or the extinct Russenorsk and Russian-Chinese Pidgin, are also known from the area. Since all of the languages mentioned in this paragraph are typically used as second languages, they are not further discussed in this presentation as it is difficult to characterise languages without native speakers as endangered. Case studies on most of these and several other similar languages are presented in the *Mixed languages* (1994). No creole languages seem to have been recorded in the area.

Also excluded from this presentation are sign languages, classical languages that represent the proto-form of a group of modern languages, notably Latin, and artificial languages such as Esperanto and Volapük, despite the fact that there may be grounds for characterising some of them as endangered.

More crucially, dialects, irrespective of their endangerment status, have not, as a rule, been included in this presentation. Contrary to the frequently repeated claims about the impossibility of distinguishing between language and dialect on a purely linguistic basis, it is maintained here that in the overwhelming majority of cases linguistic criteria lead to an unambiguous solution, and while borderline cases, such as large dialect continua, may exist, they are in reality much rarer than is often assumed

and their documentation is typically anecdotal. Notably, the linguistic concept of language is only concerned with spoken varieties, or vernaculars, and does *not* depend on their social status or level of cultivation, and the presence or absence of a corresponding literary tradition plays no role at all in defining a language. The foundation of this concept lies in the nature of language boundaries as opposed to dialect boundaries. A single language is typically a contiguous network of intergradating varieties, while two separate languages, even when closely related, are divided by a unique boundary, which has historically emerged through the recurring process of extinction of transitional dialects. Language boundaries must be corroborated by the more commonly cited criterion of mutual intelligibility to the extent that dialects of a single language must be mutually intelligible, but, obviously, there should be no requirement of total mutual unintelligibility between closely related languages. The point is that while mutual intelligibility occurs in all possible shades, rendering it useless as a criterion on its own, dialect and language boundaries represent two reasonably distinct states of affairs, which function differently with regard to historical change, areal contacts, and bilingualism or diglossia.

Needless to say, there is another, equally legitimate but different concept of language, based on sociolinguistic criteria such as literary use or administrative divisions. From a sociolinguistic point of view, it is entirely plausible to recognise Bosnian, Croatian and Serbian as separate entities following the separation of literary traditions, while from a linguistic perspective these traditions are based not only on a single language conventionally referred to as Serbo-Croat but more exactly on varieties of the same dialect group within it. Similarly, national boundaries mark Alsatian, Aranese and Torne Valley Finnish as mediums of separate language communities, and thus languages in the sociolinguistic sense, while they continue to function as parts of strictly linguistic units, Alemannic, Gascon and Finnish, respectively. Also, only five Jewish languages, Italkian, Judezmo, Juhur, Krimchak, and Yiddish, are recognised here, because they represent markedly distinct units as spoken languages in contrast with other varieties often referred to as Jewish languages but probably better described as dialects, registers or literary traditions; most notable of the spoken varieties not listed below are Shuadit (Judeo-Provençal; recently extinct) and Yevanic (Judeo-Greek; nearly extinct).

It is important to understand that all minority language communities, including those demarcated by cultural, political and administrative boundaries and thus qualifying as distinct languages in the sociolinguistic sense such as Alsatian, Aranese and Torne Valley Finnish mentioned above, or Valencian as opposed to Catalan spoken in Catalonia, Rhenish Franconian in France and Luxemburgian in Belgium to name just a few other examples, need and deserve equal support and full recognition of their language rights, irrespective of the fact that they have no specific entries in this presentation. Crucially, decisions about the use and cultivation of the native language must have the support of the local population without unwanted interference from external authorities.

The details of distinguishing between language and dialect in allegedly or factually problematic cases are presented in the following discussion on the classification of the languages dealt with in this chapter. In anticipation, it can be immediately said that the majority of the cases where handbooks typically state that the status of a particular variety is unclear derive from confusing linguistic and sociolinguistic, or synchronic and diachronic criteria. To name a few examples, there should be no doubt at

all that Frisian, Mordvin, Nenets and Saami do *not* constitute single languages but can only be understood as groups of closely related but distinct languages. Notably, even in this presentation which seeks to recognise all units defined by unambiguous language boundaries as separate languages, many languages with highly divergent dialects remain, for instance, Albanian, Armenian, Chuvash, Dargwa, Estonian, Kryts, Nanay, Nivkh, North Frisian, Ossete and Romani.

On the other hand, some concessions have been made in a number of cases where some transitionality or intergradation is found between two closely related varieties that nevertheless are generally viewed as separate languages. Such pairs of what could be called co-languages or twin languages include Bulgarian-Macedonian, Galician-Portuguese and Azerbaijani-Turkish. The point is that a shallow language boundary can be found between the members of these pairs, but the boundaries do not follow national borders but, in the above-mentioned cases, Macedonian is spoken in south-western Bulgaria, Galician in northern Portugal, and Azerbaijani in eastern Turkey.

Of course, national borders do not generally correlate with language boundaries except on the level of official terminology referring to literary traditions rather than spoken languages. For instance, Luxemburgian is spoken not only in Luxembourg but in all neighbouring countries, Belgium, France and Germany, while it is usually not called Luxemburgian but Moselle Franconian or simply Franconian in those countries. Similarly, the border between the Netherlands and Germany does not correspond to any language boundaries, but Low Saxon is spoken across the border in the Netherlands, while Dutch in the form of the Brabantish dialect extends to Germany, and the Limburgian-Ripuarian language area covers parts of both countries.

It can be argued that in a number of cases the present classification is over-differentiating, for instance Mari, Komi (including Permyak), Occitan (covering all Oc languages), French (covering all Oïl languages), Sardinian, Abkhaz (including Abaza), Circassian and Koryak (including Alutor) could basically be treated as single languages rather than groups of closely related languages, but the current solutions have simply turned out to be the most fruitful in the preparation of this presentation.

While dialects as such are excluded from this presentation, there exists a special type of varieties which still to some extent function within larger language units as dialects, but are geographically or ethnologically clearly detached from them and have started to develop independently, especially because of their different language environment. They include, for example, the Algherese dialect of Catalan, the Norman dialects of the Channel Islands, Cimbrian and other Germanic enclaves in Italy, Dolgan as a distinct ethnolinguistic variety of Yakut, or the ancient forms of Greek in Italy known as Griko. Quite obviously, such varieties need to be discussed in the context of language endangerment, and they also require a specific label, for which I here suggest 'outlying dialect'; other possible terms could be 'diaspora dialect' or 'isolated dialect', but they appear slightly restrictive.

While the list of distinct languages presented below is supposed to be reasonably comprehensive, the decision about what counts as an outlying dialect is obviously a more relative issue. Some omissions are therefore inevitable, and several examples come immediately to mind. To give an idea, Neo-Chuvan, a dialect of Russian with notable lexical influences from aboriginal Siberian, especially Yukagir languages, is not included here because it seems too integrated into Russian-speaking society, to the extent that it might be considered a specific register rather than a dialect, but it remains a borderline case nevertheless. Another example would be Tazy (in Chinese

pinyin *Dazi*), a dialect of Northern Chinese, influenced, in particular, by Udege, and spoken mainly in the village of Mikhaylovka in Ol'ga County in Maritime (Primor'ye) Region in the Russian Federation by 100 to 200 people. Tazy is even included in the *Krasnaja kniga jazykov narodov Rossii* (1994), but since it has only been isolated from the rest of Northern Chinese for decades rather than centuries, it does not yet qualify as an outlying dialect. In Europe, the list of outlying dialects could probably be augmented by several others that are no less distinct than those dealt with below, for instance, Luxemburgian in the Bistrița (Bistritz) region in Romania, or other Germanic varieties, collectively known as 'Transylvania Saxon'.

On the other hand, there are a number of languages whose origin can straightforwardly be tracked to an outlying dialect but which have later become fully-fledged languages. They include Kalmyk, Khamnigan Mongol, Manchurian Kirghiz, Sibe and Vojvodina Rusyn, which are *not* called here outlying dialects of Oyrat, Buryat, Khakas, Manchu or Eastern Slovak, respectively.

## Remarks on classification

### **Basque language**

Basque is probably the world's best-known language isolate, in other words it constitutes a family of its own. Language isolates and small language families are by no means rare in the world, but since Europe is dominated by a couple of large families, Basque may appear as a kind of an anomaly, although it simply represents a state of affairs that was much more widespread in early historical times, as witnessed by Etruscan and other attested non-Indo-European languages of Western and Southern Europe. The dialects of Basque are less distinct than occasionally suggested in popular sources, and continue to form a cohesive network typical of a single language.

### **Uralic language family**

The traditional classificatory scheme into which the Uralic (or Finno-Ugrian) languages are cast, found in practically all handbooks, is a binary or an almost binary model, reminiscent of the classification proposed for the Indo-European languages by Schleicher and still occasionally referred to, for instance, by Gamkrelidze and Ivanov (1995). Especially since the critique by Häkkinen (1984), however, it has become increasingly clear that there is little factual basis for maintaining the binary classification, and that its long-standing popularity is simply due to the conservativeness of the field of Uralic studies, at the same time when the classification of Indo-European languages has been thoroughly and critically examined, and the binary model found severely problematic. There are no less than three recent handbooks of the family, namely *The Uralic languages* (1988), the *Jazyki mira: ural'skie jazyki* (1993) and *The Uralic languages* (1998), but they can all be said to suffer from a number of shortcomings, not least in questions of classification.

In its structure, Uralic is actually very similar to the Indo-European family. It consists of nine distinct and cohesive branches: Saami, Finnic, Mordvin, Mari, Permian, Hungarian, Mansi, Khanty and Samoyed. The intermediate subgroups of the binary model, including Finno-Saami, Volgaic (Mordvin and Mari), Finno-Volgaic, Finno-Permian, Ob-Ugrian (Mansi and Khanty), Ugrian (Hungarian and Ob-Ugrian), and,

finally, Finno-Ugrian, the last one covering all languages except those of the Samoyed branch, are not based on shared innovations but their only meaningful, yet unreliable, justification derives from lexicostatistic calculations with Finnish as their starting-point. From the point of view of historical phonology, in particular, even the branches that appear closest to each other in the binary model, such as Saami and Finnic, or Mansi and Khanty, are so distant that their differentiation must have started directly from Proto-Uralic. The Samoyed branch is lexically innovative, though not to much greater extent than the other branches, but in other respects, Samoyed deserves no special position in the classification of the family. The statement of Comrie (1981: 93–4) that ‘the initial division of Uralic into Samoyedic and Finno-Ugric is a major division, and in many ways the modern Samoyedic and Finno-Ugric languages are very different from one another’ is therefore patently untrue both diachronically and synchronically. The rejection of the binary classification does not imply that there are no important similarities between branches such as Saami and Finnic, or the Ugrian branches, Hungarian, Mansi and Khanty, but they can be better explained by assuming language contacts than positing separate proto-languages, and, crucially, there are similarities of the same kind and extent across the alleged intermediate subgroups.

Another conservative feature of Uralic studies is that, following historical criteria or administrative decisions, units that are from a synchronic point of view clearly groups of closely related languages are called single languages with highly divergent dialects. Since the latter concept must be seen as an anomaly, the actual number of Uralic languages included in this presentation is, in almost every branch, notably higher than in conservative assessments. There are thus altogether forty-seven languages in the family, of which forty are living languages.

Consequently, there are no less than ten living Saami languages plus one that has become extinct. The internal diversity of the Saami branch easily matches that of, for instance, the Slavonic or Romance branches of the Indo-European family. Saami is very much a language chain, but the boundaries between languages are mostly clear, and often deep, except that in the part of the chain from Pite Saami via Lule Saami to North Saami there is admittedly some transitionality. The boundaries between Skolt Saami and Akkala Saami, on the one hand, and Akkala Saami and Kildin Saami, on the other, are also relatively shallow.

In older terminology that is better avoided today, calques such as ‘Baltic Finnic’ or ‘Balto-Finnic’ were often used for Finnic in the current sense of the branch of closely related languages including Estonian and Finnish, while ‘Finnic’ was supposed to cover all ‘Finno-Ugrian’ languages except ‘Ugrian’, and thus synonymous to ‘Finno-Permian’. To be able to give a consistent view of the Finnic branch, Ingrian, Karelian, Olonetsian and Lude must all be granted language status, in contrast to administrative solutions where Karelian, Olonetsian and Lude are subsumed under ‘Karelian’. In the northern Finnic language chain from Finnish via Karelian, Olonetsian and Lude to Veps, each pair of neighbouring languages are close to one another, yet language boundaries are clear, and Karelian, for instance, is not necessarily more similar to Olonetsian than it is to (Eastern) Finnish. Võro-Seto is here recognised as a regional language in relation to Estonian; it may also be called an incipient language (cf. Salminen 2003).

Another term reflecting the older usage was ‘Volga-Finnic’ instead of current ‘Volgaic’. This is another intermediate subgroup, uniting the Mordvin and Mari branches, but the relatively few common features shared by them can be rather easily explained

on the basis of language contacts. Both Mordvin and Mari were formerly regarded as single languages but they actually consist of two languages each.

In the Permian branch, Komi (Komi proper, or Zyryan) and Permyak are very closely related but differ in a number of diagnostic features. Understandably they constitute a subbranch as opposed to Udmurt.

Hungarian is a single language, but with an outlying dialect, Csángó Hungarian. By contrast, both Mansi and Khanty are branches consisting of four and three closely related languages, respectively. In the case of the two surviving Khanty languages, the relationship is actually more distant than, for instance, that between English and Swedish.

The history of the internal classification of the Samoyed branch reminds us of that of the whole language family in that it has usually also followed a binary model, involving a primary division into northern and southern branches. There is little foundation for such a scheme, but Samoyed is actually a chain of seven distinct branches, four of which are single languages, while Enets, Nenets and Selkup are subbranches consisting of closely related but separate languages despite their administrative treatment as single languages.

### ***Indo-European language family***

The difference in the traditions of classifying Uralic and Indo-European languages, despite the factual similarity in the structure and time depth of the families, is evident in, for example, Campbell (1998: 168–9). That the Indo-European family is currently represented by nine branches, Slavonic (Slavic), Baltic, Germanic, Celtic, Romance, Albanian, Hellenic (Greek), Armenian and Indo-Iranian, is almost uncontroversial. Slavonic and Baltic branches are often regarded as united, but this view could perhaps be challenged by claiming that the two groups of Baltic, West Baltic and East Baltic, should be elevated to the status of primary branches instead, and since West Baltic is extinct, we are in any case left with nine branches of modern languages. The term Romance, of course, corresponds to the entity that was created when the first language boundaries emerged within the proto-language of the branch, Latin.

The internal classification of the Slavonic branch is rather well-established and the individual languages are aptly presented in *The Slavonic languages* (1993). The closeness of East Slavonic languages is often exaggerated, and here, by contrast, Rusyn is recognised as a language distinct of Ukrainian. There may be a couple of regional languages awaiting recognition, for instance, Upper Silesian, included in parentheses in the checklist of languages below.

As to the classification of the (East) Baltic languages, only the recognition of Latgalian as a regional language in relation to Latvian needs to be mentioned.

The division of the Germanic languages into three primary branches appears solid, and problems are found at lower levels of classification. They revolve around the question of which varieties are to be subsumed under the German language. The solution offered here excludes the westernmost, and most distinct, Franconian varieties as well as Alemannic and Bavarian from German in the narrow sense, which then consists of the East Central German dialects, historically mainly based on Thuringian, plus East Franconian and Rhenish Franconian. A great deal of arbitrariness is inevitably involved in this simple answer to such a complex question, but it seems to be consistent with both the historical development and the synchronic functioning of the

varieties in question; cf. König (2001: 92). Outside this area, a number of regional languages are recognised but the list is not to be regarded as anything like the last word. The traditional classification of Germanic, referring more to literary traditions than spoken languages, forms the basis of the recent handbook *The Germanic languages* (1994), while *Germanische Rest- und Trümmersprachen* (1989) focuses on languages that enter the discussions more rarely, such as Cimbrian and Norn.

The extant Celtic languages offer few problems for classification, and the picture becomes extremely clear from the presentations in *The Celtic languages* (1993), which unlike most volumes of its kind, offers in-depth analyses of issues in language endangerment and language revival as well.

The classification of the Romance languages is arguably more complicated than that of any other language group, as can be seen from the discussions in *The Romance languages* (1988), Posner (1996), *The dialects of Italy* (1997), or, indeed, the vast compendium *Lexikon der Romanistischen Linguistik* (1988–2002), but Romance specialists also seem to enjoy emphasising this state of affairs quite eagerly. Haiman and Benincà (1992), for instance, discuss the position of the Raeto-Romance languages rather tendentiously, occasionally confusing early historical developments and more recent contact phenomena, which allows them to undermine the position of Raeto-Romance as a distinct branch, and similar statements have been aired on other forums about Occitano-Romance, in particular. These two branches, as well as Istro-Romance represented by the often neglected Istriot alone, are here recognised as well-founded subbranches alongside the less controversial Ibero-Romance, Gallo-Romance, Sardinian, Italo-Romance, Dalmatian and Romanian subbranches. The division of Romance into separate languages in this presentation can certainly be regarded as generous, but the language boundaries found between them turn out to be much clearer than is often assumed in discussions focusing on the relativity of the issue based on the idea of extensive dialect continua. Importantly, there is no obstacle to recognising larger units as languages in the sociolinguistic sense; for instance, Occitan is also in my view a valid concept of a literary language that largely unites the Oc languages described here. Incidentally, the classification of the Oc languages, itself a subbranch of Occitano-Romance, turns out to be less problematic than that of the Oïl languages, in which case the current treatment of Poitevin-Saintongeais, Gallo, Norman, Picard and Walloon as separate languages, but Lorrain and Champenois among the dialects of French, must be seen as highly tentative.

Albanian is here classified as a single language, although there are major differences between the Geg and Tosk dialects. Two outlying dialects, of Tosk to be precise, are found in Italy and Greece.

Rather than a single Greek language, there is a Hellenic branch consisting of (Modern) Greek and Tsakonian, and Cypriot Greek could also be seen as a distinct regional language. The outlying dialects of Greek, Griko (in Italy), Pontic Greek, Mariupolitan Greek and Cappadocian Greek, approach the status of separate languages.

Armenian is currently classified as a single language, although the differences between West Armenian and East Armenian are quite notable. No outlying dialects of Armenian are recognised here, which is possibly an omission, as at least Homshetsma, mentioned on various occasions by Vaux, could well be described as such.

Of the many languages of the Indo-Iranian branch, only those spoken in enclaves within the geographic area are covered by this chapter, namely Romani and Lomavren

belonging to the Indic subbranch, and Juhur and Tat as well as Ossete of the Iranian subbranch. The classification of Indo-Iranian languages is discussed in the Middle East and South Asian chapters.

### ***Semitic language family***

Since Malta and Cyprus are included in this chapter, two Semitic, and more exactly Arabic languages, Cypriot Arabic and Maltese, are considered here. Arabic was introduced to these islands only in the Middle Ages. Discussion on the classification of Arabic and Semitic can be found in Chapter 4.

### ***Kartvelian language family***

Few people would claim today that Georgian is *the* Kartvelian language, but the language status of Laz, Mingrelian and Svan is generally recognised. The only exception is that Laz and Mingrelian, which undoubtedly belong together historically, are occasionally called dialects of ‘Zan’. There is, however, little question about Laz and Mingrelian currently functioning as distinct languages.

### ***Abkhaz-Adyge language family***

The internal classification of the Abkhaz-Adyge family is uncontroversial in that it involves three branches, Abkhaz, Ubykh and Circassian (or Adyge). There are some grounds for regarding these three units as single languages, but the division of both Abkhaz and Circassian into two languages is also based on appropriate linguistic criteria. It must be noted, though, that the boundary between Abkhaz and Abaza does not follow the administrative border, but Abaza, as defined linguistically, only covers the Tapanta dialect group, while the other dialect spoken on the Russian side of the border belongs to Abkhaz proper.

A specific feature of this language family is that a very large portion of the speakers live outside their indigenous territory after the large-scale migration to Turkey, the Middle East and the Balkans in the nineteenth century to avoid Russian hegemony.

### ***Nakh-Daghestanian language family***

The primary division within Nakh-Daghestanian is, indeed, between Nakh and Daghestanian (Nichols 2003), although some schemes have dissolved the Daghestanian branch into a number of branches at the same level as Nakh. An open question is whether or not Dargwa and Lak form a subbranch together. The position of Kхингалуг and perhaps some other languages customarily included in the Lezgian subbranch has raised doubts as well, but the discussion of these matters must be left to better-informed commentators. The boundaries of the Nakh-Daghestanian languages are well-established both in literature and in reality. The only unorthodox decision made here is the treatment of Inkhokvari as an independent language rather than a co-dialect of Khvarshi, but the grounds for this as presented by Starostin seem convincing. On the other hand, this presentation remains on the conservative side in regarding Dargwa as a single language rather than a group of closely related languages.

### **Turkic language family**

The Turkic family is fundamentally divided into two distinct branches, known as Bulgar Turkic and Common Turkic, although this division is often not highlighted in the literature, presumably because the Bulgar Turkic branch, although it represents the first historically attested Turkic expansion from Asia to Europe, is survived by a single modern language, Chuvash. There are two recent handbooks covering all of Turkic, namely *The Turkic languages* (1998) and the *Języki mira: tjurkские языки* (1997), of which the latter in general offers more information about the less well-known Turkic languages. The shortcomings of the English-language handbook are compensated by the article 'Discoveries on the Turkic linguistic map' (<http://www.srii.org/Map.pdf>) by Johanson.

The classification of the Common Turkic branch into subbranches is a relatively complex matter, as evident from the detailed treatment by Schönig (1997–8), but in Central Asia and Europe, three major groups can be recognised as follows. First, the Oguz subbranch consists of Gagauz, Turkish, Azerbaijani, Kashkay, Khorasan Turk, Turkmen and Salar, of which Kashkay is spoken in Iran (and in a broad sense may also cover Iraqi Turk) and Salar in China, which means that they are dealt with in the Middle East and East Asian chapters. Second, the Karluk subbranch includes Uzbek and (Modern) Uygur. Third, the Kipchak subbranch incorporates three subgroups, namely those including (i) Crimean Tatar, Karaim, Kumyk and Karachay-Balkar, (ii) Tatar and Bashkir, and (iii) Nogay, Karakalpak and Kazakh.

In some classifications, Kirghiz, and occasionally Southern Altay as well, are included in Kipchak, but in the present scheme, they form the fourth subbranch of Common Turkic. Siberian Turkic in general appears to be a strictly geographic concept, and besides the one including Southern Altay with Kirghiz, three distinct subbranches of Common Turkic are attested in Siberia and adjacent regions. To the Yenisey Turkic, or Khakas, subbranch belong Northern Altay, Shor, Chulym Turk and Khakas, as well as Manchurian Kirghiz, which has clearly developed into a full language from being originally an outlying dialect of Khakas. The two other Siberian subbranches of Common Turkic are also known by their principal languages, Sayan Turkic by Tuvan and Lena Turkic by Yakut. There are further two peripheral subbranches, Khalaj in Iran, often included in Oguz but apparently distinct from it and dealt with in Chapter 4, and Saryg Yugur in China, the entry on which can be found in Chapter 5.

Altogether, there are thus nine reasonably clear-cut subbranches of Common Turkic, although the situation becomes complicated if administrative considerations are taken into account beside linguistic facts. Uzbek, for instance, contrary to the above divisions, conventionally includes dialects called Oguz Uzbek and Kipchak Uzbek. The status of these dialects, especially with respect to the neighbouring dialects of Turkmen and Kazakh, respectively, remains obscure even on the basis of the recent handbooks mentioned above. It seems therefore necessary to rely on the older but still indispensable compendium *Philologiae Turcicae Fundamenta I* (1959), in which Wurm simply regards Kipchak Uzbek as a Kazakh dialect, while Oguz Uzbek refers to Uzbek dialects that have been influenced by secondary contacts with Turkmen. Similar questions arise elsewhere as well, most notably in the cases of Crimean Tatar (see Doerfer in *Philologiae Turcicae Fundamenta I*) and Urum (see Podolsky 1985a; 1985b).

In this introduction, only the major Turkic languages have been mentioned, but the explicit classification found in the list of languages includes all of the languages and outlying dialects that belong to the Turkic family, and the less well-known of them have their own entries below.

### ***Mongolic language family***

The classification of the Mongolic family has become increasingly clear thanks to a number of recent studies, summarised in *The Mongolic languages* (2003). Essentially, there are five branches in the family, including a central or core group and four peripheral or satellite groups. The core branch consists of the principal language of the family, Mongolian, which includes Khalkha, Chakhar and Khorchin among its dialects, as well as Khamnigan Mongol, Buryat, Ordos and the Kalmyk-Oyrat language complex. Besides the core branch, one of the satellite branches, Dagur, is dealt with in this chapter. Of the other satellite branches, Shira Yugur, a single language, and Monguor, consisting of four languages, appear in Chapter 5, while the fifth branch, Mogol, has an entry in Chapter 4. In addition to the volume mentioned above, the *Jazyki mira: mongol'skie jazyki, tunguso-man'čžurskie jazyki, japonskij jazyk, korejskij jazyk* (1997) presents a survey of Mongolic, and it also covers the next language family, Tungusic.

### ***Tungusic language family***

The classification of the Tungusic family is perhaps slightly less complicated than is often tacitly assumed, at least if a binary classification is not required, as the family consists of four rather distinct branches (Janhunen 1996: 78; 2004). The first, North Tungusic branch, comprises, on the one hand, Even and, on the other, Evenki with closely related Negidal and Solon. The second branch consists of Udege and Oroch, and the third one of Nanay, Ulcha and Orok. These two branches are usually grouped together as Amur Tungusic, but there appear to be few features that could be regarded as shared innovations. The fourth branch is that of Manchu and Sibe, originally an outlying dialect of Manchu but by now rather a fully separate language, both dealt with in Chapter 5. The language boundaries within Tungusic are also somewhat clearer than could be expected on the basis of recent literature (cf. Grenoble and Whaley 2003, and the Tungusic Research Group website <http://www.dartmouth.edu/~trg/>), the main source of confusion being the idiosyncratic system of ethnonyms used in China. The details are presented in the entries on Evenki and Solon, but to explain the situation briefly, 'Ewenke' is the official Chinese term for the group of people whose traditional language is Solon, while 'Oroqen' is used for the majority of Evenki-speaking groups. On the diversity of the dialects of Nanay, known as 'Hejen' in China, see the respective entry.

### ***Nivkh language***

Nivkh is here seen as a single-language family, or language isolate, although differences between Amur and Sakhalin dialects as well as within Sakhalin dialects themselves are known to be extensive. Nevertheless, the northern Sakhalin dialects are closer to those on Amur, and in that sense, Nivkh can be seen as constituting a network

of dialects rather than a group of two languages. The *Języki mira: paleoaziatskie języki* (1997) is the best general source currently available, but an English-language handbook of Nivkh and the other Paleoasiatic language families is under preparation.

### **Yeniseian language family**

Ket is often considered to be an isolate, but it is rather the principal surviving member of what used to be a larger family. The other language that survived until recently, Yug, was previously considered a dialect of Ket, but there is little question of its being a separate language. Ket and Yug constitute, however, a distinct northern branch within the family, as opposed to the two other, more southerly branches, all languages of which became extinct long ago. There are some controversies about the classification of the Yeniseian languages, but the scheme proposed by Werner in the *Języki mira: paleoaziatskie języki* (1997) appears to be well-founded. According to it, Arin and Pumpokol constitute the second, and Assan and Kott the third branch of the family. That Assan and Kott were very close to each other is clear, rather the question is if they were distinguishable at all, but Werner has been able to identify a number of specific traits in the respective corpora. Arin and Pumpokol, on the other hand, are so different from each other that their belonging to the same branch is not immediately obvious, but a substantial if small set of common innovations supports the classification.

The Yeniseian languages are particularly interesting from the areal and typological points of view. They represent the only language family in Siberia with a fully developed and apparently ancient tone system, and they also have an extremely complicated morphological structure involving, among other things, the category of grammatical gender.

### **Yukagir language family**

Yukagir is to be understood as a family of two languages, Tundra Yukagir and Forest Yukagir, the differences between them being great enough to invalidate claims of their being dialects of a single language. There are several small corpora of material from extinct members of the family, but with certain reservations mainly concerning terminological issues, all of them can be seen representing older stages of the extant languages. Two varieties known as Chuvan and Omok, both of which became extinct long ago, are particularly often regarded as distinct languages, but the solution adopted here is to classify them as dialects of Tundra Yukagir instead.

### **Chukotkan language family**

Also called Chukchi-Koryak, the Chukotkan family is often seen as a union of only two languages, Chukchi and Koryak. Kerek, however, although it has been classified as a dialect of both Chukchi and Koryak, has a number of diagnostic features which clearly mark it as a separate language, and its superficial similarity with Chukchi can also be due to secondary contacts. Alutor, by contrast, is so close to Koryak that they could perhaps be seen as a single language complex, but the solution adopted here recognises Alutor and Koryak as two separate if closely related conglomerations of dialects. The details of the actual division remain to be sorted out, but here I follow

the position of Skorik in the *Jazyki narodov SSSR V* (1968: 235) in incorporating the distinct Palana and Karaga dialects into Alutor rather than, as suggested by Zhukova in the same volume (292), into Koryak.

### ***Kamchatkan language family***

Itelmen (Western Kamchadal) is the only surviving member of the Kamchatkan family which also comprises the extinct Southern Kamchadal and Eastern Kamchadal. The differences between the three languages are, indeed, well beyond what could be called dialectal. Notably, the ethnonym Kamchadal is also used to refer to the old Russian-speaking immigrant population of Kamchatka, which may or may not have absorbed Itelmen speakers in the past.

Furthermore, three languages of the Yupik subbranch of the Eskimo branch of the Eskimo-Aleut language family are or were spoken to the west of the Bering Strait, Sirenik and East Cape Yupik exclusively, and Central Siberian Yupik also on St Lawrence Island in Alaska, but their entries are included in Chapter 1 alongside the other Eskimo-Aleut languages. The Yupik subbranch has further two languages, Central Alaskan Yupik and Pacific Gulf Yupik, spoken exclusively in Alaska. The other subbranch of the Eskimo branch is Inuit, of which the westernmost language, Seward Peninsula Inupiaq, was formerly spoken on the Great Diomedede Island belonging to the Russian Federation; the last of the Inuit-speaking elders born on the island before the entire population was relocated on the continent seem to have died by the 1990s. Aleut, which constitutes the other branch of the family, is not autochthonous to the present Russian territory but its presence there derives from a nineteenth-century resettlement on the Commander Islands whence Aleuts were brought to Bering Island from Atka and to Copper Island from Attu. The entries on Aleut and the mixed language Copper Island Aleut are therefore also included in Chapter 1; the few remaining speakers of Siberian Aleut and Copper Island Aleut, all elderly, are concentrated in the village of Nikol'skoye on Bering Island in Kamchatka Province.

### **Assessing the levels of endangerment**

Europe differs from other parts of the world in that the number of very safe languages, spoken by majority populations of independent nation-states, is exceptionally high. The assimilative policies of the same nation-states are largely responsible for another peculiar feature of Europe, namely that almost all of the other languages spoken there are endangered, in other words, the number of safe minority languages is remarkably low. In Caucasia, by contrast, there are several indigenous languages that cannot be regarded as endangered at all, and, as a rule, the level of endangerment of the smaller languages is not particularly severe. Siberia is more like the most critical areas of the world because almost all of its native languages are endangered, and in many cases nearing extinction, although not on an equally drastic pace as in North America or Australia. In this section, the areas covered in the chapter are discussed from the point of view of deciding where to draw the line between safe and endangered languages.

Except for the bald fact that most of the minority languages in Europe are endangered, their status is quite different from what is known from other parts of the world.

In particular, the level of integration of European minority communities to mainstream societies is very high, because, typically, their only distinctive features are, indeed, the language and the related aspects of culture. It is therefore possible to assimilate more or less fully to the majority community, and escape various forms of discrimination, by adopting the majority language. Stable functional bilingualism would obviously contribute to both avoiding marginalisation and maintaining the native language, but the deeply-rooted ideologies of nation-states have largely favoured, and forced, a language shift instead. Political awareness within European minority communities is currently on a high level, but activists too rarely focus on basic issues of language maintenance at home and with peers, and direct their attention to literacy projects, media coverage, and higher education instead. From the point of view of descriptive linguistics, European minority languages are relatively well studied (cf. Pogarell 1983), and most of them have native speakers trained as linguists, but there are also many lacunae in our knowledge. Primary source materials are typically scattered, and in this study I have largely relied on secondary sources. Fortunately, there are several recent publications devoted to European languages that have proved useful. There is a wealth of up-to-date sources covering the languages of the Russian Federation or all of the Soviet Union, notably Comrie (1981), Kibrik in the *Endangered languages* (1991: 257–73), the *Jazykovaja situacija v Rossijskoj Federaciji: 1992* (1992), the *Krasnaja kniga jazykov narodov Rossii* (1994), Kazakevitch in the *Lectures on language situation* (2002), and the *Jazyki Rossijskoj Federaciji i sosednih gosudarstv* being currently published, yet the superb five-volume series *Jazyki narodov SSSR* (1966–8) continues to be valuable. On the languages of France, which used to be and still largely is one of the more problematic areas, we now have the books *Vingt-cinq communautés linguistiques de la France* (1988) and Sibille (2000), and *Languages in Britain and Ireland* (2000) covers the British Isles. As for other recent sources, the *Linguistic minorities in Central and Eastern Europe* (1998) lacks linguistic detail, while the *Minority languages today* (1981 [2nd edn 1990]), the *Encyclopedia of the languages of Europe* (1998), the *Minor languages of Europe* (2001), and the *Language death and language maintenance* (2003) provide rather uneven data on a selection of endangered languages. The *Ethnologue* (13th edn 1996), employed extensively by experts on languages of other parts of the world, rarely contains unique information about the languages of Europe or other areas dealt with in this chapter. Besides the printed publications, there are plenty of the electronic sources of various quality available, of which the *Euromosaic* and the *GeoNative* sites proved most fruitful for the present purpose.

There are thirty languages that have a dominant position in at least one European country: Albanian, Belarusian, Bulgarian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Macedonian, Maltese, Norwegian, Polish, Portuguese, Romanian, Russian, Serbo-Croat, Slovak, Slovene, Spanish, Swedish and Ukrainian. Consequently, these languages are in no way endangered, with the possible exception of Belarusian, which clearly is under pressure from Russian, but whose speaker base is nevertheless broad enough to keep it in the group of safe languages.

Several other languages meet the criteria for being safe in that, although there are other languages competing with them for dominance in their areas, they are widely spoken by all generations. Faroese and Luxemburgian, while having relatively low numbers of speakers, are clearly within this group because of their position as the

most commonly spoken languages in the respective autonomous and independent states. Catalan and Tatar, with millions of speakers, are the largest minority languages in Europe and also have a strong position in their autonomous regions. Bashkir, Basque, Chuvash, Galician, Welsh and West Frisian are spoken by slightly smaller minority nations, but appear to be supported by both communities and authorities to the extent that their position is quite stable. Low Saxon, Rusyn and Scots arguably have a weaker status in the countries where they are spoken, but seem not to qualify as endangered languages because they have millions of speakers and show a reasonably stable diglossia with dominant languages, backed up by substantial and old literary and other cultural traditions.

The last group of European languages that in this presentation have been deemed to be on the safe side also involve a diglossic situation. They represent what are often called regional languages in a specific sense, meaning that they are intimately related to the dominant language of their area and, because of the lack of an established literary tradition, they can from a sociolinguistic point of view be regarded as dialects. Linguistically, however, their language status is in most cases reasonably clear. Eastern Slovak is, obviously, related to Slovak, and Western Flemish to Dutch, while Latgalian is a regional language with regard to Latvian, Jutish to Danish, Scanian to Swedish, and Extremaduran to Spanish. German is the umbrella language, so to speak, for Alemannic and Bavarian, while Limburgian-Riparian is diglossic with both German and Dutch, Venetan and South Italian (Neapolitan) and Sicilian have the same relationship with Italian. These regional languages are supposedly not losing ground, at least not very fast, to dominant languages in their most important domain as the spoken medium among family and friends. If the area of a language extends to several countries, the differences in its status are typically sharp. For instance, Alemannic covers the extremely safe Swiss German as well as Alsatian in France, which is under strong pressure from French.

No less than fifty-five presumably non-endangered European languages have been mentioned above. This figure must be contrasted with the seventy-three languages and twenty-four outlying dialects spoken in Europe that have been classified as endangered here, as can be seen in their respective entries below; the languages known to have disappeared from Europe in the modern era are summarised in the following section on extinct languages. The level of endangerment is discussed to some detail in each entry, but some general guidelines concerning the reasons why some apparent borderline cases have been regarded as endangered languages are given here.

Irish is clearly a special case, as it is an official and widely cultivated language of an independent country, but as the first language, it is used by a tiny fraction of Irish society. Romani and Yiddish, by contrast, have millions of speakers worldwide, but they are everywhere under strong pressure from dominant languages. Emilian-Romagnol, Friulian, Ligurian, Lombard and Piedmontese, regional languages of northern Italy, also have large numbers of speakers but a language shift has been so rapid in their regions that they rather fall in the category of endangered languages. The same is true about Eastern Mari, Erzya, Komi, Moksha and Udmurt, the major Uralic languages in east-central Russia, whose status is thoroughly discussed by Lallukka (1990) and corroborated by the preliminary results of the 2002 census in the Russian Federation. North Saami, Romansh and Scottish Gaelic have some official support and a reasonably stable position at least in their core areas but assimilation is still taking place and the overall numbers of speakers are smaller. Other larger minority

languages such as Corsican, Gagauz, Kalmyk, Kashubian, Nogay and Walloon, as well as Campidanese Sardinian and Logudorese Sardinian, are also losing speakers on a scale that makes it necessary to define them as endangered.

All the other European languages undoubtedly belong to the category of endangered languages. It is, of course, important to keep in mind that even in their cases, endangered is not synonymous with dying, and I, for one, would never use the word 'moribund' in the context of language endangerment. It may be useful to point out that symbols characterising the level of endangerment of European languages in the map that I prepared for the *Atlas of the world's languages in danger of disappearing* (2nd edn 2001) do not match with the definitions given to them elsewhere in the book, a state of affairs that was beyond my control.

In Caucasia, despite the many ethnic and political conflicts, there are several indigenous languages that cannot be regarded as endangered at all, and, as a rule, the level of endangerment of the smaller languages is not particularly severe. The dominant languages in the area including Anatolia are, besides Russian, the four state languages, Armenian, Azerbaijani, Georgian and Turkish. Besides them, the twelve main indigenous languages of northern Caucasian republics, Adyge, Avar, Chechen, Dargwa, Ingush, Lak, Lezgian, Kabard-Cherkes, Karachay-Balkar, Kumyk, Ossete and Tabasaran, are maintained well by the population, and the bilingualism in Russian, in contrast with the major Uralic languages in east-central Russia having similar numbers of speakers, appears both functional and stable. The status of Abkhaz in Georgia is more problematic, and the closely related Abaza is clearly less secure than the above-mentioned larger languages of northern Caucasia, but even in their case the speakers appear loyal enough to the native languages to keep them on the safe side.

If the number of non-endangered languages in Caucasia and adjacent Anatolia is eighteen, there remain twenty-nine languages and two outlying dialects (Pontic Greek and Trukhmen) that are endangered to various extents. Of these, Mingrelian is perhaps functioning as a regional language in a diglossic relationship with Georgian, but its position seems quite weak. Nogay is another language with a larger number of speakers but it is spoken in widely scattered areas and appears to be losing ground to Russian. The smaller languages in Daghestan are relatively stable in spite of the low numbers of speakers, as only Agul, Juhur, Rutul and Tsakhur have close to 20,000 speakers, but even they are under growing pressure from dominant languages. There is little question about all the other languages, spoken by much smaller nations, belonging to the category of endangered languages.

In Central Asia and Mongolia, the number of languages is much smaller. Kazakh, Kirghiz, Mongolian, Turkmen and Uzbek are the dominant languages in the area covered by this chapter, and regional languages that appear to be relatively safe are Karakalpak, Khorasan Turk, Ordos and Uygur. This leaves only Oyrat and Dukha as well as two outlying dialects, Ili Turk and Sinkiang Dagur, in the group of endangered languages. The speaker base of Oyrat is still reasonably broad, but it has a very subordinate position in modern society, so regarding it as endangered seems justified. A useful summary of the complex ethnolinguistic situation in Sinkiang is provided by Rybatzki (1994).

In Siberia and Manchuria, by contrast, only two languages, Tuvan and Yakut, supported by reasonably large nations with a certain level of autonomy, can be regarded as non-endangered, although they are clearly under pressure from Russian

like all other Siberian languages. This means that as many as thirty-eight indigenous languages and three outlying dialects (Baraba Tatar, Dolgan and Manchurian Ölot) are classified as endangered here. In most cases, there is no question about their being endangered, and a large number of indigenous languages are not only endangered but are nearing extinction. The three co-official languages of the other autonomous republics in Siberia, especially Buryat, and to lesser extent Khakas and Southern Altay, as well as the most prominent indigenous language in Manchuria, Dagur, still have relatively large numbers of speakers, but a shift to Russian or Northern Chinese has been the constant trend, which has brought them into the category of endangered languages. Besides the publications covering all languages of the Russian Federation, there are a few others with a focus on Siberia and Manchuria, notably the *Arctic languages* (1990), Janhunen (1991; 1996), Vakhtin (1993; 2001), and the *Northern minority languages* (1997).

### Extinct languages

Recently (or possibly) extinct languages have their own entries below. In Europe, this category includes Manx and Alderney French, and in Caucasia and Anatolia, Ubykh and Cappadocian Greek. In Siberia, there are seven such languages, Kamas, Kerek, Southern Khanty, Southern Mansi, Western Mansi, Sirenik (see Chapter 1), and Yug; and in Central Asia there is one recently extinct outlying dialect, Ongkor Solon.

Another group of extinct languages comprises those that were still spoken and at least fragmentarily attested after the beginning of modern times. They are also incorporated into the checklist of languages below. In what follows they appear divided into geographic sections and in an approximate chronological order of their extinction moving back in time.

In Europe, *Dalmatian* was spoken in Krk (Veglia) on the coast of Dalmatia until the death of the last speaker in 1898; earlier it was also spoken in Zarar (Zada) and Dubrovnik (Ragusa). *Norn* was spoken in the Shetland Islands until approximately 1880; earlier it was also spoken in the Orkney Islands and adjacent mainland Scotland. *Kemi Saami* was spoken in Sodankylä and Kuolajärvi (Salla) counties in Lapland Province in Finland till the mid-nineteenth century. *Old Prussian* was spoken in Samland and Kurische Nehrung in East Prussia probably till the end of the eighteenth century. *Cornish* was spoken in Cornwall, England, until the death of the last speaker, which probably took place in 1777; Cornish is currently being revived and exists in three different versions; revived Cornish cannot be regarded as endangered as the number of users seems to be constantly growing. *Polabian* was spoken in the Elbe River basin in northeastern Germany until approximately 1750. *Gothic* was spoken in the Crimea in the Ukraine perhaps till the early eighteenth century. *Cuman* was spoken in Hungary till the early seventeenth century; earlier it was spread over a large area of Eastern Europe. *Mozarabic* was spoken in southern Spain until extinction in early modern times. Two extinct outlying dialects of endangered languages, Krevin and Slovincian, are presented in the respective entries on Vote and Kashubian.

From Caucasia, only *Lomavren* needs to be mentioned here. Here it refers to the original Indic language, closely related to Romani, that became extinct when it was replaced by the Armenian-based variety here conventionally called Boshka. The time of extinction seems to be unknown.

In Siberia, *Eastern Kamchadal* was spoken on the eastern coast of central Kamchatka and in the basin of the river Kamchatka, and may have survived till the first decades of the twentieth century. *Southern Kamchadal* was spoken in the southern part of Kamchatka, possibly mainly on the eastern coast, and became extinct by the end of the nineteenth century. *Kott* was spoken in the basin of the river Kan, an eastern tributary to the upper Yenisey, till the 1850s. *Mator* was spoken on the northern slopes of the Eastern Sayan mountains, extending from the eastern part of the Minusinsk region in the west (dialects *Mator* proper and *Taygi*) to the Baykal region in the east (*Karagas Samoyed*) till the 1840s. *Pumpokol* was spoken on the upper Ket', an eastern tributary to the upper Ob', till the early nineteenth century. *Yurats* was spoken in the tundra zone to the west of the lower Yenisey basin till the early nineteenth century. *Arin* was spoken in the upper Yenisey basin, to the north of the Minusinsk region, till the late eighteenth century. *Assan* was spoken to the east of the Kott territory perhaps till the late eighteenth century. Two extinct outlying dialects of endangered languages, *Kamas* (not *Kamas Samoyed*) and *Arman*, are presented in the respective entries on *Khakas* and *Even*.

Languages that only survived into the Middle Ages are not discussed in this presentation. To give just a couple of examples, *Cumbric*, *Khazar* and *Langobardic* would belong to this group. Needless to say, languages that disappeared even earlier, such as *Etruscan*, *Gaulish*, *Hittite*, *Hurrian* or *Oscan*, are left out as well.

### Notes on language names

Most languages spoken in the area have their established names in English, but especially in the case of the minority languages of the Russian Federation there appear variant names and spellings. Occasionally they derive from a lack of familiarity with earlier studies in these languages, but there occur truly controversial and problematic issues as well, and I try to shed light on them in this section.

Most minority languages of the Russian Federation as well as some others have had their old, often derogatory names changed in the twentieth century. Conservative linguists in other countries may still employ them, regarding them as cherished relics of scholarly tradition, but, in my view, there is really nothing quaint about them. While the modern names are recommended for current usage, it is, of course, important to recognise the old ones, because they are found in many sources, so here is a checklist of old and new names: *Cheremis* → *Mari*; *Gilyak* → *Nivkh*; *Gold* → *Nanay*; (*Western Kamchadal* → *Itelmen*; *Karagas* → *Tofa*; *Lamut* → *Even*; *Lappish* (or *Lapp*) → *Saami*; *Oyrot* → *Altay*; *Ostyak* → *Khanty*; *Ostyak Samoyed* → *Selkup*; *Tavgi* (or *Tavgi Samoyed*) → *Nganasan*; *Tungus* → *Evenki*; *Vogul* → *Mansi*; *Votyak* → *Udmurt*; *Yenisey Ostyak* → *Ket* (and *Yug*); *Yenisey Samoyed* → *Enets*; *Yurak* (or *Yurak Samoyed*) → *Nenets*; *Zyryan* → *Komi*. Obviously, most names in this list refer to groups of languages. Notice that 'Ostyak' in its widest sense referred not only to *Khanty* but also to the distantly related *Selkup* and the unrelated *Ket* (and *Yug*). Also notice that *Karagas* is still used as the name of a dialect of the extinct *Mator* language, and that *Southern Kamchadal* and *Eastern Kamchadal* are valid names for the extinct languages related to *Itelmen*.

Furthermore, many old language names contained the word 'Tatar', for instance, *Karachay-Balkar* was known as *Mountain Tatar*, *Khakas* as *Abakan Tatar* or *Yenisey Tatar*, and *Shor* as *Kuznetsk Tatar*. Beyond the Russian Federation, there are a few

notable name changes. Romani is nowadays preferred over Gypsy, and Eskimo is not used in language names, although it continues to be valid as the name of a branch.

Different spelling conventions abound in words borrowed from Russian, and this can be seen in geographic names as well, for instance the above-mentioned language names referred to the river Yenisey and the mountain range Altay, but these are also commonly spelled as Yenisei and Altai, respectively. Here the former spelling is adopted consistently, hence, for example, the language names Nanay, Nogay, Oyrat and Uygur contain a vowel plus 'y' instead of a vowel plus 'i' used in spellings like 'Nanai', etc., Azerbaijani being an established exception. This is also a feature of the transliteration system for Russian place-names used in this presentation: since there is no risk of confusion, the letter 'y' has two different functions in that it represents both the Russian 'back i' (a vowel variant) and the 'short i' (a glide), as in the Russian name of the capital of the Republic of Chechenia, Groznyy.

The letter 'y' is also part of the transliterations of the Russian vowel letters 'yu' and 'ya', hence the language names Buryat pro 'Buriat' and Koryak pro 'Koriak'. An apostrophe marks the Russian soft sign letter in regular transliterations, but it must be omitted from language names, for example Itelmen pro 'Itel'men' and Mansi pro 'Man'si'. The digraph 'gh' is avoided in language and place names alike, as there is no point in employing it in words deriving from Russian, hence the language names Adyge, Agul, Dagur, Godoberi, Khinalug, Lezgian, Negidal, Nogay, Udege, Uygur, Yug and Yukagir instead of 'Adyghé', etc., Kirghiz being an established exception. Some people would use 'x' for velar fricative but 'kh' is clearly more suitable for English terminology, including the language names Abkhaz, Akhvakh, Botlikh, Budukh, Hinukh, Tsakhur, Ubykh, Khakas, Khamnigan, Khanty, Khinalug, Khvarshi and Nivkh rather than 'Abxaz', etc. Double 'ss', typically copied from German spelling, is avoided, as it is sometimes added to the language names Cherkes, Kamas, Karagas, Khakas, Nganasan and Tabasaran yielding 'Cherkess', etc.

Suffixes that lack any function are avoided, so that, to give a few examples, the language and branch names Kabard, Mordvin, Ossete, Samoyed, Serbo-Croat, Slovene and Vote appear instead of 'Kabardian', 'Mordvinian', 'Ossetic' (or 'Ossetian'), 'Samoyedic', 'Serbo-Croatian', 'Slovenian' and 'Votic' (or 'Votian'). A special case involves the names with an Iranian-based 'i' at the end, which is not used in the language names Ili Turk, Juhur and Tat, pro 'Ili Turki', 'Juhuri' and 'Tati', but is by necessity and convention found in Azerbaijani.

Randomly nativised or otherwise manipulated versions of language names that have already been established as Russian borrowings in English are not recommended, so that, for example, the language names Evenki, Selkup, Sirenik and Tuvan are used in this presentation instead of 'Ewenki', 'Sölkup' (or 'Sölkup'), 'Sireniksi' and 'Tyvan'. A colonial spelling 'Chukchee' persists strangely in scholarly literature in place of the modern spelling, Chukchi. Other non-recommended variants based on recent Russian colloquialisms include 'Evenk' for Evenki, 'Khant' for Khanty, and 'Tof' for Tofa. The current names of the Finnic languages Ingrian and Vote should also not be replaced with 'Izhor(ian)' or 'Vod'.

In the names of a few Caucasian languages, exceptions are made to these conventions when English spellings seem to be well-established, so that 'h' and 'w' are used when Russian would suggest 'g' and 'v', respectively, hence the languages names Hinukh pro 'Ginukh', Hunzib pro 'Gunzib', and Dargwa pro 'Dargva'. Lezgian is likewise used instead of suffixless 'Lezgi'. For Saami, there is a widely used alternative spelling Sámi.

## The language checklist

The following list summarises the genetic classification of the languages dealt with in this chapter. Language families, marked in boldface, and major branches appear as separate paragraphs, and the different levels of subbranches are represented by the number of horizontal bars between languages; some subbranches have specific names followed by a colon, while some do not. Outlying dialects are separated by the + sign, and a number of other varieties that are clearly worth mentioning and in many cases languages in the sociolinguistic sense, while classified here as dialects on the basis of linguistic criteria, are included within parentheses, also separated by the + sign. The languages and outlying dialects with an entry in the next section are marked in italics, and the extinct, or in a few specially discussed cases possibly extinct, varieties are followed by an asterisk. Alternative names are given in parentheses.

There are altogether 172 entries, of which twenty-one represent outlying dialects of non-endangered languages. Besides that, there are fourteen outlying dialects of endangered languages dealt with in the entries on the languages in question and with references in their place in the alphabetic order.

### Basque

### Uralic

Saami: Western Saami: *South Saami* | *Ume Saami* | *Pite Saami* | *Lule Saami* | *North Saami* || Eastern Saami: *Inari Saami* | *Kemi Saami*\* | *Skolt Saami* | *Akkala Saami* (Babino Saami) | *Kildin Saami* | *Ter Saami*

Finnic: *Livonian* | *Võro-Seto* | Estonian | *Vote* (+ *Krevin*\*) | Finnish (+ Torne Valley Finnish + Finnmark Finnish) | *Ingrian* | *Karelian* | *Olonetsian* | *Lude* | *Veps*

Mordvin: *Erzya* | *Moksha*

Mari: *Western Mari* | *Eastern Mari*

Permian: *Udmurt* || Komi: *Permyak* + *Yazva Komi* (Eastern Permyak) | *Komi*

Hungarian + *Csángó Hungarian*

Mansi: *Northern Mansi* | *Eastern Mansi* | *Western Mansi*\* | *Southern Mansi*\*

Khanty: *Northern Khanty* | *Southern Khanty*\* | *Eastern Khanty*

Samoyed: *Nganasan* || Enets: *Tundra Enets* | *Forest Enets* || Yurats\* || Nenets: *Tundra Nenets* | *Forest Nenets* || Selkup: *Northern Selkup* | *Central Selkup* | *Southern Selkup* || *Kamas*\* (+Koybal) || *Mator*\* (+Taygi+Karagas)

### Indo-European

Slavonic: South Slavonic: Bulgarian (+ Pomak) | Macedonian || Serbo-Croat (+Forlak) + *Burgenland Croatian* + *Molise Croatian* || Slovene + *Resian Slovene* ||

West Slavonic: Lechitic: Polabian\* | *Kashubian* (+ Slovincian\*) | Polish (+ Upper Silesian) || *Sorbian* || Czech-Slovak: Czech | Slovak | Eastern Slovak | *Vojvodina Rusyn* (Bachkan) ||| East Slavonic: Rusyn (Ruthenian) (+ Lemko [Lemke] + Boyko + Hut-zulian) | Ukrainian | Belarusian (+ Polissian [Palesian]) | Russian

Baltic: West Baltic: Old Prussian\* ||| East Baltic: Lithuanian || Latvian | Latgalian

Germanic: West Germanic: Anglo-Saxon: Scots (+ Northumbrian) | English || Frisian: West Frisian | *Saterlandic* (Sater Frisian) | *North Frisian* || Low Saxon + *Plautdietsch* || Low Franconian: Dutch + Afrikaans (see Chapter 7) | Western Flemish || Central Franconian: Limburgian-Ripuarian | Luxemburgian (Moselle Franconian) || German (+ Rhenish Franconian + East Franconian) + Pennsylvania German (see Chapter 1) | *Yiddish* (Judeo-German) || Alemannic (+ Alsatian + Swabian) + *Töitschu* || Bavarian + *Cimbrian* (+ Mòcheno\*) + *Gottscheerish* ||| North Germanic (Scandinavian): West Scandinavian: Icelandic | Faroese | Norn\* | Norwegian (+ Jemtian) || East Scandinavian: *Dalecarlian* | Swedish (+ Gutnish) | Scanian | Danish | Jutish ||| Gothic\*

Celtic: Gaelic (Goidelic): *Irish* (Irish Gaelic) | *Manx*\* (Manx Gaelic) | *Scottish Gaelic* || Brythonic: Welsh | Cornish\* | *Breton*

Romance: Ibero-Romance: Portuguese | Galician (+ Fala) || *Asturian-Leonese* (+ Sanabrese + Mirandese + Cantabrian) || Extremaduran | Spanish (Castilian) (+ Llanito) | *Judezmo* (Ladino, Judeo-Spanish) | Mozarabic\* || *Aragonese* ||| Occitano-Romance: Catalan (+Valenciano) + *Algherese Catalan* || *Gascon* (+ Aranese) || Oc (Occitan): *Limousin* | *Auvergnat* | *Languedocian* | *Provençal* (+ Niçard) | *Alpine Provençal* + *Gardiol* ||| Gallo-Romance: Oïl: *Poitevin-Saintongeais* | French (+ Lorrain + Champenois + Burgundian [Morvan] + Franc-Comtois) | *Gallo* | *Norman* + *Alderney French* + *Guernsey French* + *Jersey French* (+ Sark French) | *Picard* | *Walloon* || *Francoprovençal* + *Faetar* || *Piedmontese* | *Ligurian* (+ Monégasque) | *Lombard* | *Emilian-Romagnol* | Venetan ||| Raeto-Romance: *Romansh* | *Ladin* | *Friulian* ||| Sardinian: *Logudorese Sardinian* | *Campidanese Sardinian* ||| Italo-Romance: *Corsican* + *Gallurese Sardinian* + *Sassarese Sardinian* | Italian (Tuscan + Central Italian) | *Italkian* (Judeo-Italian) || South Italian (Neopolitan + Camponian + Abrizzian + Apulian + Lucanian) | Sicilian (+ Calabrian + Salentine [Otrantian]) ||| *Istriot* ||| *Dalmatian*\* ||| Romanian: *Istro-Romanian* | *Aromanian* (Macedo-Romanian) | *Megleno-Romanian* | Romanian (Daco-Romanian)

Albanian + *Arvanite Albanian* + *Arbëreshë Albanian*

Hellenic: Greek (+ Cypriot Greek) + *Griko* (Italiot Greek) + *Pontic Greek* + *Mariupolitan Greek* (Tauro-Romaic) + *Cappadocian Greek*\* || *Tsakonian*

Armenian

Indo-Iranian (Aryan): Iranian: South-West Iranian: Tat: *Tat* | *Juhur* (Judeo-Tat) || see Chapter 4 for the other South-West Iranian languages ||| North-west Iranian (see Chapter 4) ||| South-East Iranian (see Chapter 4) ||| North-east Iranian: Ossete || Yagnob (see Chapter 4) ||| Indic (Indo-Aryan): *Romani* | Lomavren\* | Domari (see Chapter 4) ||| see Chapter 4 for the other Indic languages

## Semitic

Arabic: Maltese | *Cypriot Arabic* | see Chapters 4 and 7 for the other Arabic and Semitic languages

## Kartvelian

Georgian || *Svan* || Zan: *Mingrelian* | *Laz* (Chan)

## Abkhaz-Adyge

Abkhaz: Abaza | Abkhaz || *Ubykh*\* || Circassian (Adyge): Adyge (West Circassian) (+ Shapsug) | Kabard-Cherkes (East Circassian)

## Nakh-Daghestanian

Nakh: *Bats* || Chechen-Ingush: Chechen | Ingush

Daghestanian: Avar-Andi-Tsezic: Avar || Andi: *Andi* | *Botlikh* | *Godoberi* | *Karata* | *Akhvakh* | *Bagvalal* | *Tindi* | *Chamalal* || Tsezic: *Tsez* (Dido) | *Hinukh* | *Khvarshi* | *Inkhokvari* | *Bezhta* (Kapucha) | *Hunzib* || Lak || Dargwa (+ Kaytag + Kubachi + Megeb) || Lezgian: Lezgian | Tabasaran | *Agul* | *Rutul* | *Tsakhur* | *Kryts* (Kryz) | *Budukh* | *Archi* | *Udi* || *Khinalug*

## Turkic

### Chuvash

Common Turkic: Khalaj (see Chapter 4) ||| Oguz: *Gagauz* (+ Balkan Gagauz) | Turkish + *Crimean Turkish* + *Urum* | Azerbaijani (Azeri) | Kashkay (+ Aynallu + Iraqi Turk) + Afshar (see Chapter 4) | Khorasan Turk | Turkmen + *Trukhmen* | Salar (see Chapter 5) ||| Karluk: Uzbek + *Ili Turk* | Uygur (+ Lopnor) ||| Kipchak: *Crimean Tatar* | *Krimchak* (Judeo-Crimean Tatar) | Cuman\* (+ Armeno-Kipchak\*) | *Karaim* | Kumyk | Karachay-Balkar || Tatar (+ Siberian Tatar) + *Baraba Tatar* | Bashkir || *Nogay* + *Alabugat Tatar* + *Karagash* + *Yurt Tatar* | Karakalpak | Kazakh (+ Kipchak Uzbek) ||| Kirghiz || *Southern Altay* ||| Yenisey Turkic: *Northern Altay* | *Shor* | *Chulym Turk* | *Khakas* (+ Kamas\*) | *Manchurian Kirghiz* (Fuyü Kirghiz) ||| Saryg Yugur (see Chapter 5) ||| Sayan Turkic: Tuvan (+ Toja) | *Dukha* (Tuha) | *Tofa* ||| Lena Turkic: Yakut + *Dolgan*

## Mongolic

*Dagur* + *Sinkiang Dagur* ||| *Khamnigan Mongol* | *Buryat* || Mongolian | Ordos || *Kalmyk* | *Oyrat* + *Manchurian Ölöt* ||| Shira Yugur (see Chapter 5) ||| Monguor: Mongul (Huzhu Monguor) | Manguer (Minhe Monguor) | Bonan | Santa (see Chapter 5) ||| Mogol (see Chapter 4)

## **Tungusic**

North Tungusic: *Even* (+ *Arman*\*) || *Evenki* (+ *Orochen*) | *Negidal* || *Solon* + *Ongkor Solon*\* ||| *Udege* | *Oroch* ||| *Nanay* | *Ulcha* || *Orok* (*Uilta*) ||| Manchu: *Manchu* (+ *Jurchen*\*) | *Sibe* (see Chapter 5)

## **Nivkh**

## **Yeniseian**

*Ket* | *Yug*\* || *Arin*\* | *Pumpokol*\* || *Assan*\* | *Kott*\*

## **Yukagi**

*Tundra Yukagir* (+ *Chuvan*\* + *Omok*\*) | *Forest Yukagir*

## **Chukotkan**

*Chukchi* || *Kerek*\* || Koryak: *Alutor* (+ *Palana Koryak* + *Karaga Koryak*) | *Koryak*

## **Kamchatkan**

*Itelmen* (Western Kamchadal) | Southern Kamchadal\* | Eastern Kamchadal\*

## Alphabetical list of languages

**Agul [125]** Caucasia: southern Dagestan. Spoken in fifteen villages in Agul County and five villages in Kurakh County, amongst them Richa, Burkikhan, Tpig, Kurag, Burshag, and Fite, in the Republic of Dagestan and in a number of expatriate communities elsewhere in the Russian Federation. In 1989, nearly 18,000 speakers were reported. Despite the relatively large number of speakers of all ages, there are reasons such as the presence of more dominant neighbouring languages, notably Dargwa, Lak, Tabasaran and Lezgian, as well as the increasing influence of Russian, for regarding the language as definitely endangered. Literary experiments took place in the nineteenth century and again in recent years, and there is now one local newspaper in Agul. Until recently, Agul had no role in school curriculum but now it can be taught as a subject in local schools.

**Akhvakh [115]** Caucasia: western Dagestan. Spoken in six villages in Akhvakh County and in the villages of Ratlub, Tlyanub and Tsegob in Kakhib County in the Republic of Dagestan in the Russian Federation, as well as in the expatriate community of Akhvakh-dere in Kuba (Quba) County in Azerbaijan. There may be up to 5,000 speakers, including many children, but because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it.

**Akkala Saami (or Babino Saami) [8]** Northern Europe: Kola peninsula. Earlier spoken in the village of Babino in southern Murmansk Province in the Russian Federation, from where the speakers were forcibly translocated to Lovozero, the centre of Lovozero County. There were eight speakers in the early 1990s, all elderly, and according to a recent report, now only one speaker is known to be living. Some descendants of Akkala Saami speak Kildin Saami, but most have shifted to Russian. Nearly extinct.

**Alabugat Tatar.** See *Nogay*

**Alderney French.** See *Norman*

**Algherese Catalan** [in Algherese *alguerès*] [60] Italy: an outlying dialect of Catalan spoken by an old immigrant community deriving from the mid-fourteenth century in the town of Alghero in northwestern Sardinia. The number of speakers is probably between 20,000 and 30,000. Some children learn the language, but they may stop using it at school age. For many speakers, the influence of Italian or Logudorese Sardinian is strong. Endangered.

**Alpine Provençal [66]** In France, Alpine Provençal is spoken in the departments of Ardèche (except the north and the western border areas), Drôme (except the north), Hautes-Alpes, the northern parts of Alpes de Haute-Provence, and the southernmost parts of Isère by perhaps over 100,000 speakers, but most of them are middle-aged or elderly. In Italy, it is spoken in the upper valleys of Piedmont (Val Mairo, Val Varacho, Val d'Esturo, Entraigas, Limoun, Vinai, Pignerol, Sestriero) by approximately 100,000 speakers of all ages, but younger people are shifting to Italian. Definitely endangered. Often regarded as a single language with Provençal. – **Gardiol [67]** Italy: spoken in

Guardia Piemontese in Calabria. Gardiol is an outlying dialect of Alpine Provençal, which has been under strong South Italian influence for a long time. There are a few hundred mainly older speakers. Severely endangered.

**Alutor** [in Russian *aljutorskij jazyk*] (spelling ‘Alyutor’ better avoided) [179] Siberia: spoken in the region of the Kamchatkan Isthmus, including most of Olyutor and Karaga counties and the north of Tigil’ County in Koryak Autonomous District in the Russian Federation. Includes three dialect groups, Alutor proper, Karaga, and Palana, all of which are often considered dialects of Koryak. There are much fewer than 2,000 speakers, including very few if any children; most speakers are elderly; cf. Koryak. Kibrik reports in the *Endangered languages* (1991: 268) that of the 400 ethnic Alutor living in the village of Vyvenka in Olyutor County, ‘fewer than 100 speak the language’. Alutor dialects are still used for interethnic communication by many of the neighbouring ethnic groups in central and northern Kamchatka, including speakers of Even and Itelmen. There has never been a specific Alutor literary language, and the use of written Koryak has turned out to be impossible. Severely endangered.

**Andi** [110] Caucasia: western Daghestan. Spoken in nine villages, including Andi, Muni (Munib), and Kvankhidat’, in Botlikh County in the Republic of Daghestan in the Russian Federation. The number of speakers is usually given as being below 10,000, but the actual figure may be as high as 20,000, and includes many children. Nevertheless, for reasons such as the dominant position of Avar the language must be regarded as definitely endangered. Andi is not a written language, and Russian and Avar are used in schools.

**Aragonese** [in Aragonese *aragonés*] [59] Spain: originally spoken in most parts of the historical province of Aragón and adjacent parts of Navarra. Today Aragonese is spoken mainly in the high valleys of the Pyrenees, notably Aragon River (Somon-tano), Sobrarbe and Ribagorza, in the northern part of the Huesca (Uesca) region in the northeast of Aragón, main towns being Graus and Sabiñánigo (Samianigo). The number of active speakers is perhaps below 10,000, but there are several times more people with some knowledge of the language. Few children learn the language, but Aragonese is also actively studied as a second language. Definitely endangered.

**Arbëreshë Albanian** [96] Italy: an outlying dialect of (Tosk) Albanian deriving from the language of fifteenth- and sixteenth-century (or even earlier) refugees spoken in nearly fifty scattered villages mainly in Avellino, Potenza, Taranto, Cosenza, Catanzaro and Palermo provinces by probably under 100,000 speakers out of a much larger ethnic population. In some places, there may be children learning the language, but many stop using it at school age. Speakers’ competence is likely to vary greatly. Definitely endangered.

**Archi** [130] Caucasia: southern Daghestan. Spoken in the village of Archi (Archib) in Charoda County in the Republic of Daghestan in the Russian Federation by up to 1,000 speakers, including a number of children. Within its family, Archi is highly divergent. There is no literacy in it, and Avar and Russian are used in school and for wider communication. It has been influenced by Lak and Avar. Definitely endangered.

**Aromanian** (or *Aromunian*, or *Macedo-Romanian*) [93] the Balkans. Spoken in two larger areas, one in Thessaly and Epirus in northern Greece and another in the central parts of Albania, as well as in several small pockets in the Republic of Macedonia, Greek Macedonia and Bulgaria. In Serbia, especially between Niš and Kladovo, and elsewhere in former Yugoslavia, there are traditionally Aromanian-speaking immigrant groups deriving from the eighteenth century known by the names Vlach and Tsintsars; since the early twentieth century, there have also been many Aromanian immigrants in Romania. The total number of speakers has been estimated at 350,000, but because of a rapid language shift especially in Greece, the factual figure may be lower; on the other hand, recent data from Albania suggest up to 400,000 speakers there alone. In any case, the largest number of speakers is probably now found in Albania, followed by Greece and the Republic of Macedonia. Some children learn the language, but many stop using it at school age. Severely endangered.

**Arvanite Albanian** [95] Greece: an outlying dialect of (Tosk) Albanian spoken in old immigrant communities that occupy extensively the rural areas in Attica, Boeotia (Voioῑ́a), southern Euboea (Évvoia), northern Andros (Ándros), Corinthia, Argolis and a few smaller sections of Peloponnese and central Greece. There are perhaps 50,000 speakers, but a language shift has proceeded rapidly and regular use is probably confined to a much smaller number of people. Very few if any children learn the language. All idiolects are heavily influenced by Greek. Severely endangered.

**Astrakhan Nogay.** See *Nogay*

**Asturian-Leonese** (also known as *Bable*; in Portugal known locally as *Mirandese* [*mirandês*]) [57] Spain earlier in most parts of the historical provinces of Asturias and León, extending to the northeastern corner of Portugal. Asturian-Leonese (or simply *Asturian* [*asturianu*]) survives mainly in Asturias around Oviedo (Uviéu) where some children learn the language, and to much lesser extent in the western parts of León Province (*Leonese*) and in the Sanabria region in Zamora Province (*Sanabrese*) in the autonomous region of Castille-León, where probably all speakers are elderly. The number of speakers in Spain is not known exactly but a reasonable estimate is 100,000. Though not recognised as official, the language is to some extent used in the school system in Asturias. In Portugal, the language is spoken in the Miranda do Douro region and in two towns in the Vimioso region by some 10,000 to 15,000 people; it was officially recognised there in 1999, and it has gained some presence in education, media and public life since then. Nevertheless, Asturian-Leonese is increasingly endangered both in Portugal and Spain.

**Auvergnat** [in Auvergnat *auvernhàs* or *auvernhat*] [63] France: spoken in an area covering the departments of Cantal (except Aurillac region), Haute-Loire, and Puy-de-Dôme, and extending to the Gannat region in Allier, the Saint-Bonnet-le-Château region in Loire, and the western border areas in Ardèche. There may be over 100,000 people able to speak Auvergnat, but very few if any children learn it. Even the language of the most competent speakers is likely to be heavily influenced by French. Severely endangered.

**Bagvalal** (also called *Bagulal*) [116] Caucasia: western Daghestan. Spoken in the villages of Khushtada, Tlondoda, Kvanada and Gemerso in Tsumada County and the

villages of Tlissi and Tlibisho in Akhvakh County as well as in a number of expatriate communities elsewhere in the Republic of Dagestan in the Russian Federation. The number of speakers lies somewhere between 4,000 and 6,000. Because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it.

**Baraba Tatar [142]** Siberia. An outlying dialect of Tatar spoken in the Baraba steppes in the western parts of Novosibirsk Province in the Russian Federation. Baraba Tatar has more recently been influenced by mainstream Siberian Tatar. There are 8,000 members of the ethnic group, but only the elderly speak Baraba Tatar, while the younger generations use Tatar and Russian. Schools operated in Tatar until 1964 when they were shifted to Russian. Baraba Tatar must be regarded as severely endangered.

**Bats** (also called **Batsbi**) [in Georgian *tsova-tushi*] [110] Caucasia: northeastern Georgia. Spoken in the village of Zemo-Alvani in Akhmeta County. The number of speakers is usually given as 3,000, but active users may be far fewer. Modern Bats is heavily influenced by Georgian with most speakers, and a more traditional variety of Bats is perhaps only known by a small group of elders. There is no literacy in it, but Georgian is used in school and for wider communication.

**Bezhta** (also called **Kapucha**) [123] Caucasia: western Dagestan. Spoken in the villages of Bezhta, Tlyadal, and Khasharkhota in Tsunta County in the Republic of Dagestan in the Russian Federation. There may be up to 7,000 speakers, including many children, but because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it, but Avar and Russian are used in school and for wider communication.

**Botlikh** [112] Caucasia: western Dagestan. Spoken in the villages of Botlikh and Miarsu in Botlikh County in the Republic of Dagestan in the Russian Federation. There are approximately 4,000 speakers, including many children, but because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it, but Avar and Russian are used in school and for wider communication.

**Breton** [in Breton *brezhoneg*] [56] France: spoken mainly in western Brittany. In the 1950s, there were over a million speakers, but only 240,000 speakers were reported in 1997. The transmission of the language to children stopped almost completely between 1950 and 1970. Now a small number of children are learning the language, but it is not clear if they continue to use it in adulthood. Regular users are mostly fully competent, although French influence is felt strongly. Severely endangered.

**Budukh** [129] Caucasia: northern Azerbaijan. Spoken in the village of Budukh and in the expatriate communities of Deli-Kaya and Pirusti in Kuba (Quba) County and elsewhere in Azerbaijan. There are approximately 1,000 speakers, including some children; higher figures found in literature probably refer to members of the ethnic group. There is no literacy in it, but Azerbaijani is used in school and for wider communication. Definitely endangered.

**Burgenland Croatian [40]** Austria: an outlying dialect of Serbo-Croat spoken in Burgenland. It is said to differ extensively from Serbo-Croat dialects spoken in Croatia. There were approximately 28,000 speakers in the 1970s, now probably much less. It is possible that a number of children still learn the language, but many families have shifted to the majority language. Speakers' competence is not known exactly, but varies within different areas, and, presumably, between age groups. Definitely endangered.

**Buryat [in Chinese pinyin *Buliyate*] [159]** Siberia, China and Mongolia. Two old dialect groups, Western Buryat in Cisbaykalia, to the north of the Eastern Sayan mountains and in the Angara region, and Eastern Buryat in Transbaykalia, extending from Lake Baykal in the west to the Onon basin in the east, can be distinguished. Western Buryat is spoken mainly in Ust'-Orda (Ust'-Ordynskiy) Buryat Autonomous District and elsewhere in Irkutsk Province in the Russian Federation; an outlying dialect of Western Buryat is spoken in the Barguzin valley of Transbaykalia. Eastern Buryat is spoken in the Republic of Buryatia and in Aga (Aginskiy) Buryat Autonomous District and elsewhere in Chita Province in the Russian Federation, as well as by recent immigrant groups in adjacent regions of Mongolia and China. Old Bargut and New Bargut are earlier branches of Eastern Buryat in Manchuria deriving from the seventeenth and eighteenth centuries, respectively. Old Bargut [in Chinese pinyin *Chen Baerhu*] is spoken in the northern part and New Bargut [in Chinese pinyin *Xin Baerhu*] in the southwestern part of the Barga steppe region in Chen Baerhu Banner and Right and Left Xin Baerhu banners, respectively, in Hulun Buir League in Inner Mongolia in China. There are probably over 400,000 speakers of all dialects, including many children, but Buryat is everywhere under pressure from dominant languages, Russian, Chinese and Mongolian. In many rural areas of the Russian Federation, children are learning the language, but in cities Russian is normally spoken even between Buryat speakers. Among the Western Buryat there are relatively few child speakers, and Russian influence has been stronger on Western Buryat than on Eastern Buryat. Speakers in China and Mongolia, especially among the Old Bargut, are shifting to Mongolian. Buryat has a literary standard in Cyrillic script based on Eastern Buryat, which replaced the earlier use of Written Mongolian; Written Mongolian serves as the literary language for the Buryat population in China. In China and Mongolia all Buryat varieties are officially counted as varieties of Mongolian. Buryat as a whole is moderately endangered, the Buryat-speaking territory continually shrinking and there being very few monolinguals, but the situation of dialects such as Western Buryat is even more alarming, and Old Bargut may be regarded as severely endangered.

**Campidanese Sardinian [86]** Italy: spoken in southern Sardinia. The number of speakers is perhaps approximately 500,000; cf. Logudorese Sardinian. Many children learn the language, but often stop using it at school age. Definitely endangered.

**Cappadocian Greek [100]** an outlying dialect of Greek spoken in a few isolated communities in the interior of Cappadocia in central Turkey, notably in Sille (Silli) near Konya, villages near Kayseri, and Faras (Pharasa) and adjacent villages, before the genocide of 1915 and the subsequent population exchanges, after which most survivors settled in Greece. Cappadocian Greek is extinct in Turkey and in Greece, and the Greek dialects still spoken in the regions of Tsalka in Georgia and Alaverdi in

Armenia, which derive partly from Cappadocian Greek, have apparently assimilated to Pontic Greek. Recently extinct.

**Central Selkup [37]** Siberia: in the basin of the upper Ob' and on its tributaries, from the Chaya in the south to the Tym in the north, with each river basin having a dialect of its own, mainly in the eastern parts of Kargasok and Parabel' counties in Tomsk Province in the Russian Federation. There are a couple of hundred speakers at best, probably all elderly. There exists a movement of national awakening, but it seems mainly concerned about social and ecological problems rather than the fate of the native language. There have been attempts to create a written standard in Cyrillic script. Critically endangered.

**Chamatal** [in Russian *čamalinskij jazyk*] [118] Caucasia: western Daghestan. Spoken in the villages of Upper Gakvari, Lower Gakvari, Agvali, Tsumada, Richaganik, Gadyri, Gigat', and Kvankhi in Tsumada County in the Republic of Daghestan in the Russian Federation. The number of speakers lies somewhere between 4,000 and 6,000. Because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it, but Avar and Russian are used in school and for wider communication.

**Channel Islands French.** See *Norman*

**Chukchi** [in Russian *čukotskij jazyk*] (also known as *Luoravetlan*) [177] Siberia: in and around Chukotka from the Arctic coast to the Bering Sea region, covering most parts of Chukchi Autonomous District, adjacent regions in Koryak Autonomous District and Lower Kolyma (Nizhnekolymskiy) County in the Republic of Sakha (Yakutia) in the Russian Federation. There are two geographical and economical groups known as the Maritime or Coastal or Sea Chukchi and the Inland or Reindeer Chukchi, but there is no corresponding dialectal division. In 1989, 11,000 speakers were reported, but the figure may be slightly inflated. Chukchi is often thought to be rather safe, but there are actually very few children speaking the language. There used to be phonological, possibly also lexical and grammatical, differences between men's speech and women's speech, but the differences are now being levelled in favour of the men's variety. Chukchi has a written standard in Cyrillic script, but it is little used in education. Definitely or even severely endangered.

**Chulym Turk** (also called *Chulym Tatar* or *Middle Chulym language*) [150] Siberia: on the middle Chulym, a tributary to the upper Ob', mainly in Tegul'det County in Tomsk Province and Tyukhtet County in Krasnoyarsk Region in the Russian Federation. Two extinct Turkic varieties known by the special names Kūärik and Kecik can apparently be regarded as dialects of Chulym Turk. The number of fluent speakers is only thirty-five; all of them are middle-aged or older. There are a few others with strong interference from Russian and Siberian Tatar, which is spoken on lower Chulym. Critically endangered.

**Cimbrian** (or *Cimbro*) [in German *Zimbrisch*] [51] Italy: an outlying dialect of Bavarian spoken in the towns of Giazza (Ljetzan) in Verona Province, Roana (Rowan), Mezzaselva (Mitterballe) and Rotzo (Rotz) in Vicenza Province, and Luserna (Lusern)

in Trento Province. Earlier spoken in two larger areas known as *Sette comuni* (Sieben Gemeinden) established in the twelfth century, including Roana and Rotzo as well as Lavarone (Lafraun) from which Luserna was founded in the sixteenth century, and *Tredici comuni* (Dreizehn Gemeinden) established in the thirteenth century, of which only in Giazza the language has survived up till today. In Luserna, Cimbrian is still the language used by the whole community of some 500 people, while the couple of hundred speakers in Giazza, and the approximately 1,500 speakers in *Sette comuni* are mostly middle-aged or older. Although a number of children apparently learn Cimbrian in Luserna, the language cannot be regarded as safe even there. Speakers are all fluent in Italian and Venetan, many knowing German as well, and modern Cimbrian tends to be heavily influenced by these languages. Although Cimbrian has some presence in school and media, and there have been attempts to promote it more actively in recent decades, it is definitely endangered. Another outlying dialect of Bavarian known as *Mòcheno* has been spoken in the Fersina Valley (Fersental, Valle del Fersina, or Val dei Mòcheni) in Trento Province. Although there are reports that *Mòcheno* still has 400 speakers in Fierozzo (Florutz), 1,000 in Palù (Palai), and 460 in Frassilongo (Gereut), it seems more likely that they currently use a local variety of Venetan with numerous Germanic substate influences, and at least Matzel in *Germanische Rest- und Trümmersprachen* (1989) refers to *Mòcheno* as extinct.

**Corsican [87]** France: spoken by approximately two thirds of the total population of 250,000 people of the island of Corsica (Corse). Also spoken on Maddalena Island off the northeast coast of Sardinia. There are tens of thousands of speakers in expatriate communities elsewhere in France and other parts of the world. The position of Corsican is much stronger than that of any other minority language in France, but it is nevertheless increasingly endangered. – **Gallurese Sardinian [88]** Italy: spoken in northeastern Sardinia. Gallurese Sardinian is an outlying dialect of Corsican. The number of speakers is perhaps approximately 100,000; cf. Logudorese Sardinian. Many children learn the language, but often stop using it at school age. Definitely endangered. – **Sassarese Sardinian [89]** Italy: spoken in northwestern Sardinia. Sassarese Sardinian is here treated as another outlying dialect of Corsican. It has been influenced by other Romance languages, notably Logudorese Sardinian, Ligurian and Tuscan, so profoundly that it can also be called a mixed dialect. The number of speakers is perhaps approximately 100,000; cf. Logudorese Sardinian. Many children learn the language, but often stop using it at school age. Definitely endangered.

**Crimean Tatar [139]** Eastern Europe. Originally spoken in the Crimea, but most speakers were deported from there mainly to Central Asia after the Second World War. There are also Crimean Tatar speakers in the southern parts of the Dobruja region in northeastern Bulgaria and adjacent Romania. Crimean Tatar was spoken across the central parts of the Crimea, while speakers of Nogay in the northern steppe area and those of Crimean Turkish in the southern coastal region have been officially subsumed under the Crimean Tatar. In 1989, 250,000 speakers were reported from the Soviet Union, including over 40,000 in the Ukraine, where they lived both in the traditional area in the Crimea and elsewhere, especially in Kherson Province, 180,000 in Uzbekistan plus a few from Tajikistan, and 20,000 in the Russian Federation, mostly settled in Krasnodar Region; since then, many more have returned to the Crimea. It is unlikely, however, that these figures represent the number of fully fluent speakers reliably:

Comrie writes in the *Atlas of the world's languages* (1994) that ‘Maria S. Polinsky advises me that the number of solid first language speakers of Crimean Tatar may not exceed 100,000’. In Bulgaria, the official term ‘Tatar’ refers principally to Crimean Tatar, and the number of ‘Tatar’ speakers lies somewhere between 6,000 and 11,000. In Romania there are up to 25,000 ‘Tatar’ speakers, which covers speakers of both Crimean Tatar and Nogay. In some communities, children learn the language, but many stop using it in adulthood. Probably all speakers exhibit strong influence of related Turkic languages, Russian or Bulgarian, and younger people may be less competent in Crimean Tatar. Definitely endangered.

**Crimean Turkish [135]** Eastern Europe: an outlying dialect of Turkish originally spoken in a number of villages in the Yalta region along the southern shores of the Crimea, but the speakers shared the fate of the Crimean Tatars, as explained above. There are a few thousand speakers at best, but even this rough figure is based on a best guess rather than a reliable estimate. Definitely endangered.

**Csángó Hungarian [23]** Eastern Europe: an outlying dialect of Hungarian spoken in the Moldova region in Romania, primarily in Bacău, Neamţ and Iaşi counties; formerly, there were also a few Csángó-speaking villages east of the river Prut in the Republic of Moldova. The number of speakers is in the range of some tens of thousands. Csángó Hungarian derives from the Middle Ages; there are also Hungarian speakers in Moldova who are not Csángó but are mainly seventeenth- and eighteenth-century immigrants from the Szekler community in Transylvania. There are still young people able to speak Csángó Hungarian in a few villages, but they rarely use it amongst themselves. The attitude of Romanian authorities towards the language continues to be hostile. Severely endangered.

**Cypriot Arabic** (also called *Cypriot Maronite Arabic*) [105] Cyprus: spoken mainly in the Kormakiti village in the north. There are far fewer than 2,000 speakers out of many thousand Maronites. Probably no children learn the language, and younger speakers, if there are any left, are more fluent in Greek. Greek influence has been strong for a long time to make Cypriot Arabic a hybrid language. Many people, especially the young, have moved to the government-controlled, Greek-speaking area from the self-declared Turkish Republic of Northern Cyprus. Severely endangered.

**Dagur** (also spelled *Daur*) [in Chinese pinyin *Dawoer*] [156] Manchuria: spoken in central and northwestern Manchuria in China, mainly in Morin Dawa Daur and Ewenki autonomous banners in Hulun Buir League in Inner Mongolia (in the Nonni and Imin basins) and in Heilongjiang Province (in the Nonni basin) as well as within a small population in the Aihui region (in the middle Amur basin); cf. Janhunen in the *Northern minority languages* (1997: 128–9). In earlier times the Dagur language area covered parts of adjacent Siberia. There are probably more than 50,000 speakers. In Hulun Buir (especially in the Imin basin) the language is being retained well, while it has ceased to be transmitted to children among most groups in Heilongjiang, where (as in the cities of Nenjiang and Qiqihar) most or all speakers are middle-aged or older. Speakers are generally fluent within those communities retaining the language. Many attempts have been made to write Dagur in various systems (Manchu, Mongolian, Roman, Cyrillic), and a project of a new literary language is currently being

planned. Increasingly endangered. – *Sinkiang Dagur* (also called *Turkestan Dagur*) Sinkiang: spoken in the Ili (Yili) region of Sinkiang, in Tacheng Prefecture in Xinjiang Uygur Autonomous Region in China. Sinkiang Dagur is an outlying dialect of Dagur. Dagur speakers were translocated to Sinkiang from Manchuria in 1763. The number of speakers is not known exactly but it lies in the range of a few thousand. Sinkiang Dagur is possibly being learnt by some children. Speakers are in many cases reportedly fluent but under the influence of neighbouring languages. There is widespread multilingualism in Kazakh, Uygur, Chinese and other languages. Definitely endangered.

**Dalecarlian [157]** Northern Europe: Sweden. Spoken in the central parts of Dalarna Province by approximately 10,000 people. Its most characteristic dialect (occasionally referred to as ‘Elfdalian’) is spoken in the south of Älvdalen County by 3,000 speakers. In Älvdalen, Dalecarlian is maintained quite well and it has recently created literary standard, while in other areas younger people are shifting to Swedish/ Because of the small size of the community the language must be regarded as definitely endangered.

**Dolgan [155]** Siberia: southern Taymyr peninsula. Spoken mainly in Dudinka and Khatanga counties in Taymyr (Dolgan and Nenets) Autonomous District as well as in Anabar County in the Republic of Sakha (Yakutia) in the Russian Federation. Dolgan is an outlying dialect of Yakut with deep influences from Evenki. In 1989, 5,500 speakers were reported, and it is one of the few smaller languages in Siberia with a substantial number of child speakers. Nevertheless, because of the small size of the community and the dominant position of Russian, Dolgan must be regarded as definitely endangered.

**Dukha** (also spelled **Tuha**) [in Mongolian known as *tsaatan*] **[153]** Mongolia: spoken by a semi-nomadic group living to the northwest of Lake Khövsgöl, mainly in the Tsaaganuur district of Khövsgöl League in the northernmost region of Mongolia, in an area bordering the Republic of Tuva in the west and the Republic of Buryatia in the northeast. Dukha is closely related to but clearly distinct from Tuvan, although this is not always made clear in general handbooks. The number of speakers lies between 100 and 200, and because of the geographic isolation and the traditional way of life, practically all of the members of the group continue to speak the native language, and at least until the 1970s children used to be monolingual in it. Given the very small population, Dukha must nevertheless be regarded as definitely endangered.

**Eastern Khanty [30]** Siberia: spoken along the western and eastern tributaries to the middle Ob’, from the Vas’yugan to the Pim, mainly in Nefteyugansk, Surgut and Nizhnevartovsk counties in Khanty-Mansi Autonomous District and the adjacent parts of Tomsk Province in the Russian Federation. Eastern Khanty dialects include Salym Khanty, Surgut Khanty and Vakh-Vas’yugan Khanty. There are probably a couple of thousand speakers, including some children, but many of them shift to Russian at school age; most active speakers are therefore middle-aged or older. Definitely endangered.

**Eastern Mansi** (also called **Konda Mansi**) **[25]** Siberia: in the basin of the river Konda, a western tributary to the lower Irtysh, mainly in Konda County and the southern

parts of Sovetskiy County in Khanty-Mansi Autonomous District in the Russian Federation. The number of speakers is not known exactly, but most likely there are only a handful of elderly speakers, and some reports are even more pessimistic. Nearly extinct.

**Eastern Mari** (often called **Meadow Mari** after the largest dialect group) [20] Eastern Europe: east-central Russia. Spoken mainly in central and eastern parts of the Republic of Mariy-El and parts of Bashkortostan, Tatarstan, the Udmurt Republic, and Yekaterinburg, Perm' and Orenburg provinces, in the Russian Federation. In 1989, 540,000 speakers were reported for both Mari languages, the large majority of them, approximately 500,000, speaking Eastern Mari. Many children learn the language, but only those living in remote rural areas continue to use it actively; in eastern diaspora areas in Bashkortostan practically all children learn the language. In many publications, the term 'Eastern Mari' is reserved to the diaspora groups outside the Republic. Increasingly endangered.

**Emilian-Romagnol** [81] Italy: spoken in the region of Emilia-Romagna, parts of the provinces of Pavia, Voghera, and Mantua in southern Lombardy, the Lunigiana district in northwestern Tuscany, the Republic of San Marino, and the Pesaro-Urbino province in the region of Marche. The number of speakers is probably between one and two million but the domains of use of Emilian-Romagnol have been constantly narrowing. Increasingly endangered. Emilian and Romagnol can also be viewed as two separate languages.

**Erzya** [17] Eastern Europe: Central Russia. Spoken mainly in the eastern parts but also in small pockets in the northwest and south of the Republic of Mordovia in the Russian Federation; also several areas in Nizhniy Novgorod, Samara, Saratov and Orenburg provinces, Tatarstan and Bashkortostan. In 1989, 770,000 speakers were reported for the two Mordvin languages; approximately 500,000 of them were speakers of Erzya. Quite a lot of children learn the language, but most of them use it only with elderly relatives, which means that they will have little use for it in the future. The situation is actually better in the Eastern diaspora areas than in the Republic. Increasingly endangered.

**Even** [163] Siberia: in northeastern Siberia, from the Lena to the Anadyr', including areas in Allaykhovskiy, Momskiy, Tomponskiy, Middle Kolyma (Srednekolymskiy), Verkhoyansk, Ust'-Yansk, Kobay, Lower Kolyma (Nizhnekolymskiy), Upper Kolyma (Verkhnekolymskiy), Oymyakonskiy, Abyy, Bulun and Eveno-Batantay counties in the Republic of Sakha (Yakutia), Ola, Northern Even (Severo-Evenskiy), Omsukchan, Ten'kinskiy and Yagodnoye counties in Magadan Province, Okhota County in Khabarovsk Region and Bilibino and Anadyr counties in Chukchi Autonomous District in the Russian Federation, plus a small diaspora group in the Bystraya region of central Kamchatka in Kamchatka Province and several areas in Koryak Autonomous District. In 1989, 8,000 speakers were reported, but the figure is probably slightly inflated, so that the approximate number of speakers may be given as 7,000. There are very few or no child speakers; most speakers are middle-aged or older; the whole population is bi- or multilingual in Russian, Yakut or Koryak. Two orthographical standards, both in Cyrillic script, conforming to Yakut and Russian,

respectively, have limited use. Severely endangered. – *Arman* was an archaic outlying dialect of Even spoken in a small coastal pocket at the Okhotsk Sea, to the south west of Magadan city, that became extinct recently, though no exact date is available.

**Evenki [164]** Siberia (including parts of the Russian Far East), Manchuria and Mongolia. Evenki is the widest-spread language of Siberia, spoken by a population sparsely covering the whole taiga zone from the Yenisey in the west to the lower Amur and Sakhalin in the east, and from Taymyr and the lower Lena in the north to Baykal and the upper Amur in the south where the language area extends to the Khingan region of northern Manchuria. In the Russian Federation, the largest numbers of Evenki communities are found in the Republic of Sakha (Yakutia), Evenki Autonomous District including adjacent regions of Taymyr (Dolgan and Nenets) Autonomous District and Krasnoyarsk Region and Khabarovsk Region including settlements in Maritime (Primor'ye) Region and Sakhalin Province, but the level of native language maintenance is higher in the remote parts of Amur and Chita provinces, as explained by Atkine in the *Northern minority languages* (1997); Evenki is also spoken, in declining numbers, in a few localities in Irkutsk and Tomsk provinces and the Republic of Buryatia. In the Russian Federation, 10,000 speakers were reported in 1989. The majority of them are middle-aged or older but Evenki is still learnt by some of the children in some localities, notably in the Tunguska region in Evenki Autonomous District and in the middle Amur region. In official Chinese terminology, 'Evenki' [in Chinese pinyin *Ewenke*] refers almost exclusively to the closely related but distinct Solon, while all long-established varieties of Evenki are subsumed under 'Orochen' [in Chinese pinyin *Elunchun*, in nativised pinyin *Oroqen*]. The official Oroqen nationality has 8,196 ethnic members, but Evenki is only spoken by at most 30 per cent of the group, mainly middle-aged to elderly. They represent four dialects of Evenki: the remaining speakers of *Orochen* proper, all over fifty years old, live scattered around Oroqen Autonomous Banner in Hulun Buir League in northeastern Inner Mongolia, while the speakers of the other dialects live along the Amur basin in northwestern Heilongjiang Province: *Kumarchen* (or *Manegir*), in the Kumara basin on the upper Amur, still has some child speakers, but the closely-related *Selpechen*, in the Fa and Selpe basins on the middle Amur, has no speakers under fifty, and *Birarchen* (or *Birar*), in the Xun and Zhan basins, is only spoken by a small number of elderly people. Only two recent immigrant groups of Evenki speakers belong to the official Ewenke nationality: the Manchurian Reindeer Tungus (also known as 'Yakut Ewenke'), numbering less than 1,000 people, who migrated from Russia to Chen Barag Banner in Hulun Buir League in Inner Mongolia in the early nineteenth century, and the Khamnigan (also known as Tungus Ewenke) who are originally from the Onon-Argun region of Transbaykalia on the Siberian and Mongolian sides of the border but who settled mainly in Chen Baerhu Banner in Hulun Buir League after the October Revolution in 1917; today, there are approximately 1,000 Evenki-speaking Khamnigan who also speak Khamnigan Mongol as a second language. Altogether, there are a couple of thousand speakers of Evenki in China, including very few children except among the Khamnigan where Evenki is learnt consistently as the home language by most children. In spite of the vast area, the dialectal differences within Evenki are small, and the case for regarding Evenki and Orochen, and possibly also Birarchen and Kumarchen, as separate languages is weak; notably, some Evenki-speaking groups living to the north of the Amur are also occasionally referred to as Orochen.

Not only Russian and Northern Chinese but also more prominent indigenous languages, notably Yakut and Buryat, are spoken as a second, or often the first, language in Evenki communities. Evenki has a written norm in Cyrillic script, used for elementary-school textbooks and occasional other publications; the language has been taught in China at a few elementary schools in IPA transcription, but with poor results. Severely endangered in the Russian Federation and China, possibly extinct in Mongolia.

*Faetar*. See *Francoprovençal*

**Forest Enets** (also called *Bay*) [33] Siberia: spoken in the forest zone on the lower Yenisey, now concentrated in the village of Potapovo in Dudinka County in Taymyr (Dolgan and Nenets) Autonomous District in the Russian Federation. In the last 150 years, the speakers of Forest Enets have gradually moved approximately 500 kilometres northwards along the Yenisey basin. The number of speakers is currently perhaps between 30 and 40, and practically all of them are middle-aged or older. The last speakers are mainly trilingual, speaking also Russian and Tundra Nenets. Critically endangered.

**Forest Nenets** [36] Siberia: spoken in Pur County and the southern parts of Nadym County in Yamal Nenets Autonomous District and the adjacent parts of Nizhnevartovsk, Surgut and Beloyarskiy counties in Khanty-Mansi Autonomous District in the Russian Federation. There are approximately 1,500 speakers; cf. Tundra Nenets. A few children learn the language, but many shift to Russian at school age. In some areas there is notable Khanty influence, and younger people may prefer Russian. Definitely endangered.

**Forest Yukagir** (also called *Southern Yukagir* or *Kolyma Yukagir* or *Odul*) [176] Siberia: spoken in the forest zone, on the sources of the Kolyma, divided between the Republic of Sakha (Yakutia) and Magadan Province in the Russian Federation, now largely concentrated in the village of Nelemnoye as well as in the county centre Zyr-yanka in Upper Kolyma (Verkhnekolymskiy) County in the Republic of Sakha, previously in a much wider area in the upper Kolyma region. There are less than fifty speakers, all middle-aged and older, fluent speakers are found among those over sixty years old; cf. Tundra Yukagir. Speakers are typically multilingual. Of the two Yukagir languages, Forest Yukagir seems to be dying even more rapidly than Tundra Yukagir. Critically endangered.

**Francoprovençal** [76] In France, most or all speakers of Francoprovençal are elderly, but they are still found in almost all departments within the traditional area of the language: in Savoy (Savoie and Haute-Savoie), Francoprovençal is spoken by approximately 35,000 speakers, in Ain (mainly in the region of Bresse) by 15,000, in Rhône by 1,000, in Loire by 5,000, in the northern and central parts of Isère by 2,000, and in the southern parts of the departments of Jura and Doubs by 2,000; the original area extended to the northernmost parts of Ardèche and Drôme as well. In Italy, Francoprovençal (locally called *Harpitan*) is spoken in the Aosta Valley and in the Alpine valleys to the north and east of Val di Susa in Piedmont by perhaps 70,000 speakers of all generations but with a notable shift to Italian among younger people.

In Switzerland, Francoprovençal was earlier spoken everywhere in Suisse romande except the Canton of Jura where the Franc-Comtois dialect of French was spoken, but it survives mainly in mountain villages of Valais and Fribourg, perhaps also Vaud, being most actively used in the village of Evolène in Valais. Definitely endangered in Italy; severely endangered in France and Switzerland. – **Faetar [77]** Italy: spoken at Faeto and Celle San Vito in Foggia Province in Apulia Region. Faetar is an outlying dialect of Francoprovençal, which has been under strong South Italian influence for a long time. Approximately 700 speakers were reported in 1995, probably including a few children, as the population remains loyal to the language. Because of the small size of the community and the continuing emigration, Faetar is definitely endangered.

**Friulian** [in Friulian *furlan*] [84] Italy: spoken in the Autonomous Region Friuli-Venezia Giulia except Trieste Province and western and eastern border regions, and in Portogruaro area in Venezia Province in Veneto Region. The estimates about the number of speakers range from 350,000 to 500,000 or as many as 720,000. Many children learn the language, but often stop using it at school age, especially in towns where a shift to Venetan is evident. Most idiolects are heavily influenced by Venetan and Italian. Increasingly endangered.

**Gagauz [134]** Eastern Europe. Spoken primarily in a relatively compact area known as Bujak, deriving from eighteenth- and nineteenth-century immigration, covering the towns and surrounding areas of Comrat, Ceadir-Lunga, Basarabeasca, Taraclia and Vulcănești in the south of the Republic of Moldova, constituting the Gagauz autonomous area, and the adjacent Izmail region of Odessa (Odesa) Province in the Ukraine. A separate group known as Maritime Gagauz occupies the original Gagauz territory, covering the coastal region to the north of, and inland from, Varna in Bulgaria. Scattered Gagauz settlements are or were found in other parts of Bulgaria as well as in Romania, Serbia and Central Asia. There are further at least three groups, collectively known as **Balkan Gagauz** (often subsumed under ‘Balkan Turkic’), whose linguistic connections with Gagauz proper remain somewhat obscure, namely (i) *Macedonian Gagauz* in southeastern Macedonia, apparently both within the Republic of Macedonia and in the northern parts of the Province of Macedonia in Greece; (ii) *Surguch* in the region of Edirne (Adrianople) in Turkey; and (iii) *Gajal* in the region of Deli Orman in Bulgaria. In 1989, 140,000 speakers were reported from the Republic of Moldova and 25,000 speakers from the Ukraine; the number of Maritime Gagauz in Bulgaria was reported as 12,000 in 1990 but more recent estimates indicate only 5,000 speakers. Balkan Gagauz has probably few speakers now: early in the twentieth century, there were approximately 4,000 Macedonian Gagauz, approximately 7,000 Surguch, and an unknown number of Gajal; the total number of Gagauz speakers is close to 200,000. Some children learn the language, but in many areas, the communities are shifting to dominant languages; there are probably no children speaking Balkan Gagauz. Many speakers have been strongly influenced by neighbouring languages, both by closely related Turkish and non-related languages like Russian, Bulgarian, Romanian and Greek. Increasingly endangered.

**Gallo [69]** France: spoken in the eastern parts of Brittany and in the department of Loire-Atlantique (Pays de Loire region) by a small portion of the population. Very

few if any children learn the language. Even the language of the most competent speakers is likely to be heavily influenced by French. Severely endangered.

**Gallurese Sardinian.** See *Corsican*

**Gardiol.** See *Alpine Provençal*

**Gascon [61]** In France, Gascon is spoken in the departments of Landes, Gers and Hautes-Pyrénées, the eastern parts of Pyrénées-Atlantiques, the western parts of Haute-Garonne and Ariège, and the southern and western parts of Gironde, by perhaps 250,000 speakers, and it is still used actively by many people especially in the Béarn region (Pyrénées). Nevertheless, everywhere in France, as stated by Field in FEL IV (2000: 86), it ‘has nearly ceased to be used for the primary socialization of children’. In Spain, Gascon (known as *Aranese*) is spoken by most of the 4,800 people living in the Aran Valley in the Pyrenees. At least elderly speakers both in Béarn and Aran appear to be fully competent; elsewhere in France, French influence is increasingly strong, and this may be so for many speakers in Béarn, too; in Aran, there is notable Catalan and Spanish influence. Definitely endangered in Spain; severely endangered in France.

**Godoberi [113]** Caucasia: western Daghestan. Spoken in the large village of Godoberi and two small villages Zibirkhali and Beledi in Botlikh County in the Republic of Daghestan in the Russian Federation. The number of speakers lies somewhere between 2,500 and 4,000, with relatively few child speakers. Because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it, but Avar and Russian are used in school and for wider communication.

**Gottscheerish** (also called **Granish**) [52] Slovenia: originally spoken in the Gottschee (Kočevje) region in northern Slovenia, but the speakers were resettled during the Second World War, and now live scattered practically all around the world. Gottscheerish is an outlying dialect of Bavarian that was flourishing in Gottschee but is now only known by the oldest members of the community in exile. Critically endangered.

**Griko** (also called **Italiot Greek**) [97] Italy: spoken in the Salento peninsula in Lecce Province in southern Apulia and in a few villages near Reggio di Calabria in southern Calabria. Griko is an outlying dialect of Greek largely deriving from Byzantine times. The Salentine dialect is still used relatively widely, and there may be a few child speakers, but a shift to South Italian has proceeded rapidly, and active speakers tend to be over fifty years old. The Calabrian dialect is only used more actively in the village of Gaddhiciano, but even there the youngest speakers are over thirty years old. The number of speakers lies in the range of 20,000. South Italian influence has been strong for a long time. Severely endangered.

**Guernsey French.** See *Norman*

**Hinukh** (also spelled **Ginukh**) [120] Caucasia: western Daghestan. Spoken in the village of Ginukh in Tsunta County, and by an immigrant community in the village of Monastyrskoye in Kizlyar County, in the Republic of Daghestan in the Russian

Federation. The number of speakers is somewhere between 200 and 500, with few child speakers. There is no literacy in it, but Avar and Russian are used in school and for wider communication. Hinukh is also influenced by the closely related Tsez language. Definitely endangered.

**Hunzib** (also spelled **Gunzib**) [124] Caucasia: western Daghestan. Spoken in the villages of Gunzib, Nakhada and Garbutl' (Gorbutl') in Tsunta County in the Republic of Daghestan in the Russian Federation. The number of speakers is somewhere between 600 and 800, with few child speakers. There is no literacy in it, but Avar and Russian are used in school and for wider communication. Hinukh is also influenced by the closely related Bezhta language. Definitely endangered.

**Ili Turk** [138] Sinkiang: spoken in the eastern parts of the Ili Valley in Xinjiang Uygur Autonomous Region in China. Ili Turk is an outlying dialect of Uzbek with deep influences from other Turkic languages. It is the language of thirty or so families, where younger generations are shifting to Kazakh or Uygur. Ili Turk is not officially recognised, and its speakers tend to be fluent in Kazakh and Uygur. Definitely endangered. For detailed information, see Zhào and Hahn 1989.

**Inari Saami** [6] Northern Europe: Lapland. Spoken in central Inari County of Lapland Province in Finland. There are approximately 300 speakers. Very few children used to learn the language; before the founding of the first Inari Saami language nest in 1997, the number of child speakers was down to five or so. Since then, however, the programme, which at the moment runs two language nests, has proved remarkably successful, and a substantial portion of Inari Saami youth are now growing as functional bilinguals who may even use the native language not only with elderly relatives or language nest tutors but also amongst themselves. This positive trend is fortunately supported to some extent by recent changes in the local school system. At the same time, a number of adults who have not learnt Inari Saami in childhood have started to study it actively in evening classes. Given the small size of the community and the overwhelming presence of Finnish-language media, the situation is still precarious, but clearly, most Inari Saami people value their native language highly and are committed to its survival. Despite the positive trends, it must be said that the language is being heavily influenced by Finnish, and continues to be severely endangered.

**Ingrian** [in Russian *ižorskij jazyk*, in Estonian *isuri keel*, in Finnish *inkeroinen*] [13] Northern Europe: northwestern Russia. Spoken in three small areas on the southern shore of the Gulf of Finland in St Petersburg (Leningrad) Province in the Russian Federation: the cape of Kovashi (Hevaa) in Lomonosov County, the cape of Soykin (Soikkola) in Kingisepp County, and the cape of Kurkola in the lower Luga River area in the same county; a fourth area was formerly found along the Oredezh River in Gatchina County. In 1989, 300 speakers were reported; the actual number is lower, and most of them prefer other languages. There are probably no younger speakers, so practically all speakers are middle-aged or older. Most idiolects are heavily influenced by Finnish and Russian. Severely endangered.

**Inkhokvari** (also spelled **Inkhokari**) [122] Caucasia: western Daghestan. Spoken in the villages of Kvantlyada, Santlyada, Inkhokari (Inkhokvari) and Khvayni in Tsumada

County in the Republic of Daghestan in the Russian Federation. The number of speakers is somewhere around 1,000. Inkhokvari is customarily, but with little foundation, regarded as a single language with Khvarshi. There are child speakers, but because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it.

**Irish** (also known as **Irish Gaelic**) [53] the Republic of Ireland: four principal areas in the west, two in County Donegal, one each in Galway and Kerry counties, plus eight small pockets, also in Mayo, Cork and Waterford counties; formerly also in Northern Ireland. There were 29,000 first language speakers in the four principal areas, plus less than a thousand in each of the pockets in 1976; perhaps less than 20,000 today; the official census figures include many English speakers who have learned Irish at school. A number of children learn the language, but their number appears to be decreasing in the Irish-speaking areas. At the same time, Irish is being used widely as a second language in all parts of the Republic of Ireland as well as in Northern Ireland. Shannon in FEL III (2000: 109–16) explores the place of Irish in schools, while Hindley's (1990) controversial treatment may be offered as an alternative approach. Definitely endangered in the Republic of Ireland; extinct as a first language in Northern Ireland.

**Istriot** (also called **Istro-Romance**) [91] Croatia spoken in the southwest of the Istrian peninsula, mainly in the towns of Rovinj (Rovigno), Bale (Valle), Vodnjan (Dignano) and Galižana (Gallesano). Jahn in FEL II (1998: 48) estimates that there are 400 first-language and 400 second-language speakers in Istria, plus 500 others living outside Istria. There are young speakers in Bale and middle-aged speakers in Galižana, but elsewhere speakers are elderly. All idiolects tend to be heavily influenced by Venetan, but Istriot is also actively cultivated in the form of folkloristic poetry. Severely endangered.

**Istro-Romanian** (also spelled **Istrio-Romanian**) [92] Croatia: spoken in the northeast of the Istrian peninsula, in the village of Žejane to the north of Učka mountain, and a few villages to the south of it, the main one of these being Sušnjevica. Jahn in FEL II (1998: 48) gives an estimate of 300 first-language and 100 second-language speakers in Istria, plus 1,000 others living outside Istria. Other sources contain highly inflated figures, such as 450 to 500 for Žejane and 800 to 1,000 for the southern villages. There are no children and few younger speakers. All speakers are bilingual in Serbo-Croat, which has had a strong influence on the language. Severely endangered.

**Italkian** (or **Judeo-Italian**) [90] Italy, mainly in urban areas in Rome and central and northern Italy; also in Corfu, Greece (**Corfiote Italkian**). There appear to be some speakers in Rome, but generally very few in Italy are fluent in the language. In Corfu, although the total population counts less than 50 individuals, Italkian is still in some sort of use. Critically endangered.

**Itelmen** [181] Siberia: in a small pocket on the western coast of central Kamchatka, mainly in the villages of Kovran and Upper Khayryuzovo in the south of Tigil' County in the southern corner of Koryak Autonomous District in the Russian Federation. In 1989, 500 speakers were reported, but currently there are perhaps less than 100 speakers, all elderly; the figure of the 1989 Soviet census suggesting 2,500 ethnic

Itelmen is highly inflated. There is noticeable areal influence of Palana Koryak. Critically endangered.

**Jersey French.** See **Norman**

**Judezmo** (or **Judeo-Spanish**, or **Ladino**) [58] since the Middle Ages, Judezmo, the traditional language of Sephardic Jews, has been spoken in dispersed communities mainly in Greece and Turkey, primarily in the historical provinces of Macedonia and Thrace, but also elsewhere in the Balkans, as well as in North Africa, especially Morocco (where known as **Haketia**). It is now spoken in a few locations in Turkey, where speakers are largely concentrated in Istanbul (traditionally in the quarters of Balat and Hasköy), by probably less than 10,000 speakers, in resettled communities in Israel where the majority of the speakers live, and in the United States and several European countries. In Greece or elsewhere in the Balkans there are very few if any Judezmo speakers left after the Holocaust, and in Morocco the language is extinct. According to Hetzer in the *Minor languages of Europe* (2001: 144), ‘one can say that less than 400,000 people still have a certain command of’ the language worldwide, but ‘most of them [are] older than 50 years’. Other estimates suggest a total of some 200,000 speakers, and Judezmo is not the dominant language for most speakers. Younger speakers are very few. Severely endangered.

**Juhur** (also called **Judeo-Tat**) [102] Caucasia: northern Azerbaijan and parts of Daghestan and other regions of northern Caucasia. Still spoken, though in diminishing numbers, in its most original area in Azerbaijan, mainly in the villages of Krasnaya Sloboda in Kuba (Quba) region and Vartashen (Vardashen, currently Oguz) in Vartashen (Vardashen) region, but the great majority of speakers live in communities that started to form in the eighteenth century, especially in the town of Derbend and surrounding rural areas in the Republic of Daghestan in the Russian Federation, but also in other towns in Daghestan such as Makhachkala, Majalis, Pyatigorsk and Buynaksk and elsewhere in northern Caucasia, in particular in Nal’chik in the Kabard-Balkar Republic and, until recently, in Groznyy in the Republic of Chechnya; there are some speakers in Baku in Azerbaijan as well. Juhur is customarily, but with little foundation, regarded as a single language with Tat. While there are apparently Christian Tats speaking Tat rather than Juhur in Daghestan, most members of the Tat nationality in the Russian Federation appear to only nominally distinct from the Mountain Jew nationality, whose traditional language is Juhur. In 1989, the joint figure of reported native language speakers from the Russian Federation was nearly 25,000, of them 15,000 from Daghestan, to which are added 5,000 speakers belonging to the Mountain Jew nationality from Azerbaijan. Since then, however, thousands of Juhur speakers have emigrated to Israel, where they live in places like Sderot, Haderah and Or Akiva. A Juhur literary language based on Derbend dialect is in use in Daghestan. In Azerbaijan, Azerbaijani is used as a literary language. Definitely endangered.

**Kalmyk** (also spelled **Kalmuck** or **Kalmuk**) [160] Eastern Europe: southeastern Russia. Spoken in the Kalmyk Republic and the adjacent part of Astrakhan, Rostov and Volgograd provinces and Stavropol’ Region in the Russian Federation. Kalmyk may be historically seen as an outlying dialect of Oyrat, but now it clearly functions as a

distinct language, although the differences are largely limited to lexicon. In 1989, 150,000 speakers were reported. A few children learn the language, but most of them tend to prefer Russian. Definitely endangered.

**Kamas** (or **Kamas-Koybal**) [151a] Siberia: formerly spoken in the eastern part of the Minusinsk region, in modern Krasnoyarsk Region in the Russian Federation, where the area of Kamas became gradually restricted to the Kan and Mana river basins and their sources on the northwesternmost slopes of the Eastern Sayan mountains. The last Kamas-speaking community lived here in a single village, Abalakovo. The last speaker, Klavdia Plotnikova, died in September 1989. She was trilingual in Khakas and Russian, with rather rudimentary native language skills. Recently extinct.

**Karagash.** See *Nogay*

**Karaim** [141] Eastern Europe. Originally spoken in the Crimea in the Ukraine, where small communities of ethnic Karaim remain near Yevpatoriya; since medieval times Karaim has been translocated in the west Ukrainian towns of Luts'k (Lutsk) and Halych (Galich) and in Trakai (Troki) and a few other places in Lithuania. In 1989, 500 speakers were reported, including 200 in Lithuania, and 100 in the Crimea, but in reality there are about fifty speakers in Lithuania plus a small number who have settled in Poland, only six in Halich, and none in the Crimea. Two thirds of the speakers are over sixty years old, but recent revitalisation efforts in the Lithuanian community give some hope to the language, and while severely endangered, it is not immediately nearing extinction, as described by Csátó in the *Minor languages of Europe* (2001).

**Karata** [114] Caucasia: western Daghestan. Spoken in nine villages, including Karata and Tokita (Tukita), in Akhvakh County in the Republic of Daghestan in the Russian Federation. There are at least 5,000 speakers, but because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it.

**Karelian** [in Karelian and Finnish *karjala*] [14] Northern and Eastern Europe: north-western and west-central Russia. Spoken in several separate areas in the Russian Federation: northern and central regions of the Republic of Karelia, mainly in Loukhi, Kem', Kalevala and Muyezer'skiy counties, the western parts of Belomorsk, Segezha and Medvezh'yegorsk counties, the southwestern part of Kondopoga County and the Porosozero region of Suoyarvi (Suojärvi) County; small pockets in the vicinity of Tikhvin and Novgorod; in Tver' Province, mainly around Tolmachi, Maksatikha and Ves'yegonsk. In Finland, Karelian is spoken by people evacuated, in 1940 and 1944, from former Finnish territories north of Lake Ladoga including Suojärvi, Suis-tamo and Korpiselkä counties, currently the central parts of Suoyarvi (Suojärvi) County in the Republic of Karelia, and in two border villages in Suomussalmi County in Oulu Province. In the Russian Federation, there are possibly approximately 35,000 speakers; the combined figure for Karelian, Olonetsian, and Lude in 1989 was 63,000 which may actually be too low, because many speakers outside the Republic may not have been registered as such. There are in any case more Karelian speakers in Tver' Province than in the Republic of Karelia. A number of children learn the language, but most if not all of them become more fluent in Russian and largely stop using

Karelian later in life. Following the initiative of the local cultural society (in Finnish, Uhtua-Seura) and the participation of a Finnish linguist (Annika Pasanen), two language nests have recently, in 1999 and 2002, started to operate in the northwestern county centre of Kalevala (Uhtua). A language shift in the county centre has proceeded rapidly, so that people under thirty years of age possess little knowledge of Karelian, those between 30 and 70 generally have a passive command but prefer Russian in most contexts, and only those seventy years and older use the native language in all its traditional functions. The situation is, however, much better in smaller villages of the county such as Vuokkiniemi and Jyskyjärvi, where speakers of all ages can be found. The trend is everywhere such that most Karelian speakers change to Russian when addressing young people, but there are also dedicated individuals who use Karelian as much as possible. Since the language nests have not yet produced any clear changes to the better, the choice of home language remains the decisive factor in the survival of Karelian. Speaking Karelian still bears a major stigma, so it is not surprising that shifting to Russian is seen as the easy and perhaps inevitable solution. Since the value of the native language is also felt strongly, a dilemma arises, and people escape to sweeping the problem of language maintenance under the carpet, and end up in a state of collective self-betrayal where the language is claimed to be in active use in families while everybody knows that this is not the case. In Finland, there are a few thousand speakers, all elderly; even those who are fluent in Karelian use mostly Finnish. A literary language existed in Tver' Province in the 1930s, and it has been revived recently. Another literary variant, based on the northern dialects, is now used in the Republic of Karelia. Definitely endangered in the Russian Federation; critically endangered in Finland.

***Kashubian*** (also spelled *Cassubian*) [43] Poland: dispersed in an area northwest, west and southwest of Gdansk, mainly in Wejherowo, Lebork, Bytów (Betowo), Pock (Puck), Kartuzy, Koscierzyna and Chojnice districts. The number of speakers is not known exactly, but may be around 50,000, including quite a lot of child speakers in some areas. Definitely endangered. – An outlying dialect called *Slovincian* was spoken in the parishes of Schmolsin and Garde in Pomerania in present-day Poland, until extinction in approximately 1900.

***Kerek*** [178] Siberia: originally spoken in a large belt along the Bering Sea coast between the Olyutor Bay and the Anadyr Bay, later only in the villages of Meynopil'gino and Khatyrka in Beringovskiy County in Chukchi Autonomous District in the Russian Federation. There were three elderly speakers in 1991; in the early 1950s there were approximately 100 speakers. The language of the last speakers showed strong interference from Chukchi, the language which the Kerek community had adopted, as well as from Russian. Possibly extinct.

***Ket*** (also called *Northern Ket* or *Imbat Ket*; cf. Yug) [173] Siberia: spoken on the middle and upper Yenisey and its tributaries, mainly between the Yeloguy and Turukhan basins, today concentrated in the villages of Farkovo, Serkovo, Kellog, Kangatovo, Baklanikha, Vereshchagino, Surgutikha, Verkhneimbatsk, Bakhta, Goroshikha and Maduyka in Turukhansk County in the north of Krasnoyarsk Region and, along the Podkamennaya Tunguska, in the village of Sulomay in Baykit County in the west of Evenki Autonomous District in the Russian Federation. In 1989, over 500 speakers were reported, which is confirmed by some specialists, for instance Vajda (2003: 394)

mentions 600 or so remaining native speakers, although Georg in the *Language death and language maintenance* (2003) provides a much more pessimistic evaluation. There are generally very few child speakers, but the situation varies depending on locality; in any case, most speakers are middle-aged or older. Severely endangered.

**Khakas [151]** Siberia: in the western half of the Minusinsk steppe region on the upper Yenisey in southern Siberia, mainly in the Republic of Khakasia but also in Uzhur and Sharypovo counties in Krasnoyarsk Region and in adjacent parts of the Republic of Tuva in the Russian Federation. Khakas is a conglomeration of several closely related dialects: Kacha, Sagay, Kyzyl, Koybal, Beltir and Shor. The Koybal dialect of Khakas is spoken by the descendants of the speakers of the Koybal dialect of Kamas, and the Shor dialect of Khakas derives from the migration (in the sixteenth to nineteenth centuries) of part of the speakers of Mrassu Shor across the Kuznetskiy Alatau to the Minusinsk region. In 1989, 60,000 speakers of Khakas were reported. In some rural areas, the language is being learnt by children, but in cities Russian is typically the only language even in families with Khakas-speaking parents; most speakers are therefore middle-aged or older, and Khakas has become the language of a small minority, approximately 10 per cent, of the total population of Khakasia. Nevertheless, it has recently been introduced as the language of instruction in a few schools, including one in Abakan, the local capital. Khakas has a written standard in Cyrillic script with some use even in newspapers and belletristics. Definitely endangered. – **Kamas** was an outlying dialect of Khakas spoken by the last speakers of Kamas (Samoyed) and their descendants until recently when it as well became extinct.

**Khamnigan Mongol** [in Chinese pinyin *Hamunikan*] [158] Manchuria, earlier and more originally in the Onon-Argun region of Transbaykalia in adjacent Siberia and Mongolia. Khamnigan Mongol may be historically seen as an archaic outlying dialect of Buryat, but is better classified as a distinct language, which furthermore exists in a complex symbiotic relationship with Evenki. It is currently spoken mainly in Chen Baerhu Banner in Hulun Buir League in Inner Mongolia in China by approximately 1,500 people. Khamnigan Mongol is learnt consistently by almost all children as either the first language (in monolingual families) or as the community language (in families bilingual in Khamnigan Mongol and Evenki). Although vigorous for the moment, the survival of Khamnigan Mongol in the long run is threatened by the increasing influx of Han Chinese settlers. Both Chinese and Mongolian are also present through radio and television as well as printed material. Khamnigan Mongol is spoken as the community language by all speakers, but grammatical and lexical interference from Modern Written Mongolian is occasionally present in the speech of educated people. Modern Written Mongolian is used as the literary language. Definitely endangered in Manchuria, possibly extinct in Mongolia and Siberia.

**Khinalug [132]** Caucasia: northern Azerbaijan. Spoken in the village of Khinalug in Kuba (Quba) County in Azerbaijan. The number of speakers lies somewhere between 1,000 and 2,000, and includes some children. Within its family, Khinalug is highly divergent. There is no literacy in it, but Azerbaijani is used in school and for wider communication. Definitely endangered.

**Khvarshi [121]** Caucasia: western Daghestan. Spoken in the village of Khvarshi in Tsumada County in the Republic of Daghestan in the Russian Federation by a few

hundred speakers. Khvarshi is customarily, but with little foundation, regarded as a single language with Inkhokvari. There are child speakers, but because of the very small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it.

***Kildin Saami*** [9] Northern Europe: Kola peninsula. Earlier spoken in many locations in the eastern parts of Kola County and the western parts of Lovozero County in central Murmansk Province in the Russian Federation, from which they were forcibly concentrated to the county centre Lovozero. In 1989, combined 800 speakers were reported for the four Saami languages on Kola peninsula, the vast majority of them being Kildin Saami speakers. Today there are perhaps 650 speakers, amongst them probably no children. There are some younger speakers, but most are middle-aged or older. Kildin Saami has had small-scale use as a literary language, first in the 1930s, and again in the 1990s. Severely endangered.

***Komi*** (or, to make a distinction from the Komi subbranch comprising both Komi and Permyak, ***Zyryan Komi***; the Russian-based term ‘Komi-Zyryan’ better avoided) [22] Eastern Europe: northeastern Russia. Spoken in the Komi Republic and parts of Nenets Autonomous District, Arkhangel’sk Province, Yamal Nenets Autonomous District, Murmansk Province and other regions in the Russian Federation. In 1989, 240,000 speakers were reported. Quite a lot of children learn the language, but many of them stop using Komi later in life. Increasingly endangered. Often considered a single language with Permyak.

***Koryak*** (also known as ***Nymylan***) [180] Siberia: in northern Kamchatka, from the Okhotsk Sea to the Bering Sea, extending to the upper Anadyr basin in the north, mainly in Penzhinskiy County in Koryak Autonomous District in the Russian Federation. In 1989, a combined figure 5,000 speakers was reported for Koryak and Alutor; a rough estimate would give that slightly more than one half of them, or approximately 3,000 people, speak Koryak. There are very few child speakers, although the situation may vary depending on locality; there are some younger speakers but most are middle-aged or older. Koryak has a written standard in Cyrillic script, but education is only conducted in Russian. Often considered a single language with Alutor. Severely endangered.

***Krimchak*** (also spelled ***Krymchak***; also called ***Judeo-Crimean Tatar***) [140] the Crimea: speakers lived originally among Crimean Tatar speakers, with Simferopol as the main centre. More than two thirds of the population were murdered by Germans during the Second World War. There are now 1,200 ethnic Krimchaks in the Crimea, and 600 elsewhere. In 1989, 500 speakers were reported from the Soviet Union, including 100 in the Crimea, but the figures seem meaningless, as only people born in the 1930s or earlier appear to retain fluency in Krimchak; they number perhaps 200, and even they use Krimchak rarely. Critically endangered.

***Kryts*** (also spelled ***Kryz***; also called ***Jek***) [128] Caucasia: northern Azerbaijan. Spoken in the villages of Kryz, Jek, Khaput, Yergyuj and Alyk in Kuba (Quba) County and in expatriate communities elsewhere in Azerbaijan. The number of speakers lies somewhere between 5,000 and 8,000, and includes many children, but decreasingly so.

There is no literacy in it, but Azerbaijani is used in school and for wider communication. Definitely endangered.

**Ladin** (also called *Dolomitic Ladin*) [83] Italy: spoken in several towns and villages in the Dolomites, including Badia and Marebbe in the Badia valley and Gardena in the Gardena valley in Bolzano Province (South Tyrol), Fassa in the Fassa valley in Trento Province, and Livinallongo in the Cordevole valley in Belluno Province. The number of speakers lies in the range of 20,000. Some children learn the language, but many stop using it at school age; besides, the total number of speakers is low. Definitely endangered.

**Languedocian** [in Languedocian *lengadocian*] [64] France: spoken in an area from Bordeaux in the northwest to Montpellier in the southeast, and from Toulouse in the southwest to Rodez in the northeast, covering the departments of Aveyron, Lot, Lot-et-Garonne, Tarn-et-Garonne, Tarn, Aude, Hérault, the eastern parts of Haute-Garonne and Ariège, the southern parts of Dordogne, the Aurillac region of Cantal, the western parts of Gard, and smaller areas in Lozère, Pyrénées-Orientales, and Gironde. There may well be over a million people able to speak Languedocian, but very few children learn the language, which makes the language severely endangered despite the large number of speakers and the continuing cultivation of the Occitan literary language which is essentially based on Languedocian.

**Laz** (or *Chan*) [108] Turkey and Georgia. Spoken along the Black Sea coast in the northeast of Turkey and the southwestern corner of Georgia, including the towns of Pazar (Atina), Ardeşen, Çamlıhemşin and Fındıklı in Rize Province and Arhavi (Arkabi/Arxave), Hopa (Xopa), Borçka and Sarp (Sarpi) in Artvin Province in Turkey; Sarpi is partly in the Republic of Ajaria on the Georgian side; there are also Laz villages, founded by refugees of the 1877–8 war, in the western parts of Turkey mainly in Sakarya, Kocaeli and Bolu provinces. The number of speakers in Turkey is 20,000 to 30,000, in Georgia 1,000 to 2,000, and in an expatriate community in Germany approximately 1,000. A language shift to Turkish and Georgian has proceeded increasingly rapidly. There is no literacy in Laz. Definitely endangered.

**Ligurian** [79] in Italy, Ligurian is spoken in Liguria and adjacent areas of Piedmont, Emilia and Tuscany by well over a million speakers, but the shift to Italian has been even more marked than in other parts of northern Italy. Ligurian, under the name *Monégasque*, is also the traditional language of Monaco, where it was reportedly spoken by 5,000 people, presumably second language speakers, in 1988, after being regarded as nearly extinct in the 1970s, and its area extended to the eastern corner of the department of Alpes-Maritimes in France until the twentieth century. Ligurian is still used in the town of Bonifacio in Corsica by 300–400 mainly elderly speakers, and more actively in the towns of Carloforte on the San Pietro island and Calasetta on the Sant'Antioco island off the southwest coast of Sardinia. Increasingly endangered.

**Limousin** [in Limousin *lemosin*] [62] France: spoken in the departments of Corrèze and Haute-Vienne, most of Creuse, the northern parts of Dordogne, and the eastern parts of Charente. There may well be over 100,000 people able to speak Limousin, but very few if any children learn it. In the northern dialects there are many inherent French

features, but in all areas, even the most fluent speakers exhibit strong interference from French. Severely endangered.

**Livonian [11]** Latvia: along the northern coast of Curonia in the northwest, but also scattered elsewhere; formerly also in the historical province of Livonia east of the Gulf of Riga. In 1989, fifty speakers were reported from Latvia and about the same number in parts of the Russian Federation, but these figures were clearly inflated. There is possibly only one speaker with full native competence, which makes the language nearly extinct. At the same time, there are several younger descendants of Livonian speakers who have learnt Livonian as second language and cultivate it actively.

**Logudorese Sardinian** (also called **Logudorese-Nuorese**) [85] Italy: central Sardinia. The number of speakers is perhaps approximately 500,000; the total number of Sardinian speakers appears to be slightly over 1,000,000, although many of them use Italian more often. Many children learn the language, but often stop using it at school age. Definitely endangered.

**Lombard [80]** In Italy, Lombard is spoken in the region of Lombardy (except the southernmost border areas) and in the Novara province in Piedmont, as well as in twelfth- and thirteenth-century immigrant settlements in the south of Italy and Sicily. There are still several million speakers, but particularly in towns and among younger generations the use of Lombard has diminished drastically. In Switzerland, Lombard is spoken in Ticino Canton and in the Mesolcina District and two districts south of St Moritz in Graubünden (Grigioni) by perhaps 300,000 speakers. Increasingly endangered.

**Lude** [in Lude *lüüdi*, in Russian *ljudikovskij jazyk*, in Finnish *lyydi*] [15] Northern Europe: northwestern Russia. Spoken along a narrow strip stretching from north to south in the Republic of Karelia in the Russian Federation, covering the central parts of Kondopoga County, the eastern parts of Pryazha County, the northernmost corner of Onega (Prionezhskiy) County and the Mikhaylovskoye (Kuujärvi) region in the east of Olonets County. There are a couple of thousand speakers at best, including very few children. Severely endangered.

**Lule Saami [4]** Northern Europe: Lapland. Spoken in Jokkmokk County and parts of Gällivare and other adjacent counties in Norrbotten Province (Lule Lappmark) in Sweden and in the Tysfjord region in northern Nordland Province in Norway. The number of speakers lies somewhere between 1,000 and 2,000. A small number of children learn the language, and while the trend seemed highly alarming a few years ago, the use of Lule Saami has recently been activated to some extent. Younger speakers may nevertheless prefer Swedish or Norwegian. Definitely endangered.

**Manchurian Kirghiz** (also called **Fuyü Kirghiz**) [in Chinese pinyin *Jierjisi* or *Fuyu Keerkezi*] [152] Manchuria: still spoken on the eastern bank of the lower Nonni, in Fuyu County in Heilongjiang Province in China; extinct since the early twentieth century in the Imin region of Hulun Buir League in Inner Mongolia. Manchurian Kirghiz may be historically seen as an outlying dialect of Khakas, but must now be regarded as a distinct language. There are less than ten elderly speakers left. For several

generations, the speakers of Manchurian Kirghiz seem to have been bilingual in Manchurian Ölöť; knowledge of Dagur has also been common, but today all the local languages are being replaced by Northern Chinese. Nearly extinct.

**Manchurian Ölöť.** See *Oyrat*

**Manx** (also known as *Manx Gaelic*) [54] Western Europe: originally spoken in the Isle of Man. The last speaker, Ned Maddrell, died in 1974. Native Manx people, though monolingual in English, regard themselves as a separate people. There are people living in the Isle of Man who have now learned Manx as a second language. Extinct as a first language.

**Mariupolitan Greek** (or *Crimean Greek*; also called *Tauro-Romaic*) [in Russian *tavro-rumeskij jazyk*, in Mariupolitan Greek *runéka*] [99] Eastern Europe: an outlying dialect of Greek originally spoken in the south of the Crimea, from where its speakers moved to the shores of the Azovan Sea in the Ukraine in the 1770s, and founded the city of Mariupol' (Zhdanov) and several villages in what is now Donetsk (Donetsk) Province. In 1989, 20,000 speakers of Greek were reported from the Ukraine. Mariupolitan Greek is now influenced by Russian and Ukrainian, but earlier also by Crimean Tatar. Definitely endangered.

**Megleno-Romanian** (also called *Meglenitic*) [94] Greece and the Republic of Macedonia: originally spoken in the area where the Vardar (Axios) River crosses the Macedonian-Greek border northwest of Salonika. According to a count by Atanasov (1989), there were 5,000 speakers, only 70 per cent of whom, however, remained in the area. Many speakers were deported to Turkey during the Balkan population exchanges, and some others emigrated to Romania where they established a few villages. There are probably few if any child speakers. Severely endangered.

**Mingrelian** (or *Megrel*) [in Mingrelian *margali*] [107] Georgia: western lowlands. The number of speakers is usually given as 500,000. The closely related Georgian is generally used as a literary language, and the position of Mingrelian among younger generations is precarious. Increasingly endangered.

**Moksha** [18] Eastern Europe: central Russia. Spoken mainly in the western parts of the Republic of Mordovia in the Russian Federation, and in the adjacent parts of Ryazan' and Penza provinces; also in pockets in Tatarstan and Orenburg Province. In 1989, there were approximately 250,000 speakers; cf. Erzya. Quite a lot of children learn the language, but the typical pattern is that they only use it with elderly relatives. Increasingly endangered.

**Molise Croatian** [42] Italy: an outlying dialect of Serbo-Croat deriving from the language of fifteenth- and sixteenth-century refugees from Dalmatia spoken mainly in the villages of Montemitro (Mundimitar), San Felice del Molise (Filić, or Stifilić) and Acquaviva Collecroce (Kruč, or *Živa voda*) in Campobasso Province in southern Molise. There are over 1,000 speakers among the approximately 2,300 inhabitants of the three villages, plus a number in expatriate communities; the figure of 3,500 speakers reported by Vincent in *The world's major languages* (1987: 282) seems

exaggerated. There are a few younger speakers in Montemitro, while in Acquaviva and especially in San Felice most speakers are middle-aged or older. Severely endangered.

**Nanay** (called *Hejen* [in Chinese pinyin *Hezhe*] in China) [170] Siberia (the Russian Far East) and Manchuria: in the middle and lower Amur basin, in Nanay, Amursk, Komsomol'sk, Solnechnyy, Ul'cha and Khabarovsk counties in Khabarovsk Region and Pozharskoye and Ol'ga counties in Maritime (Primor'ye) Region as well as in Poronaysk County in Sakhalin Province in the Russian Federation, and in the lower Sungari and Ussuri basins in Tongjiang, Fuyuan and Raohe counties in the northeast of Heilongjiang Province in China. Nanay comprises three dialect groups: besides Nanay proper (with many subdialects, one of which was formerly known as 'Sama-gir'), there are *Kur-Urmi Nanay* (or *Kile*, or *Kili*, or *Kiler*), a special northern group in the region of the Kur and Urmi rivers, to the north of the middle Amur basin, with Evenki influences in its phonology, and *Sungari Nanay* (or *Kilen*; includes Ussuri Nanay and Upper Amur Nanay, also called *Akani*), which is connected with Udege and spoken in China and by the small Nanay-speaking population on the river Bikin in Pozharskoye County on the Russian side. In 1989, nearly 6,000 speakers were reported from the Russian Federation, but the actual figure is much smaller, and practically all speakers are middle-aged or older. There are perhaps 3,000 speakers of Nanay proper, and a hundred or so speakers of Kur-Urmi Nanay. There are fewer than forty elderly speakers concentrated in three villages in Tongjiang and Raohe counties in China, all now over sixty years old, out of 4,640 ethnic group members; about half speak Sungari Nanay and half speak Nanay proper. There is a written standard in Cyrillic script used for native language teaching in schools in the Russian Federation, but the language of instruction is Russian. Severely endangered in the Russian Federation, critically endangered in China.

**Negidal** (also called *Ilkan Beye* or *El'kan Beye* or *Elkembey*) [165] Siberia: in the Amgun' basin, to the west of the lower Amur, in a number of villages including Tyr, Beloklinka and Kal'ma in Ul'cha County and Vladimirovka in the County of Polina Osipenko in Khabarovsk Region in the Russian Federation. The number of speakers is well below 200. Practically all speakers are elderly, and only a few of them appear fully fluent in all stylistic varieties of the language. Critically endangered.

**Nganasan** [31] Siberia: the northernmost language of the Eurasian continent; originally spoken on central Taymyr, in the regions of the Pyasina and Taymyra river systems, now largely concentrated in the villages of Ust'-Avam and Volochanka in Dudinka County and Novaya in Khatanga County in Taymyr (Dolgan and Nenets) Autonomous District in the Russian Federation; there is, however, a group of approximately 100 people who continue to lead a semi-nomadic life in the region of the river Dudypa near Ust'-Avam. In 1989, 1,100 speakers were reported, but the actual number is much smaller, perhaps only 500, which includes very few children except in the semi-nomadic group, whose members have, at least until recently, been using Nganasan more actively. As an example of cruel destiny playing havoc in a small, self-sustainable group such as this, they have lost the special breed of reindeer necessary for survival in their severe environment, and may therefore be forced to change their way of life even in the case that Russian authorities do not meddle with them. Outside this group, only people aged forty or more are fully fluent. There are a

few old speakers with little knowledge of Russian, or only with a knowledge of a special Russian-based Taymyr pidgin also known as Govorka. Middle-aged and younger speakers are fully bilingual in Russian, with interference features easy to detect because of the intricate system of morphophonological alternations typical of traditional Nganasan. Some knowledge of Dolgan is also common. Severely endangered.

**Nivkh [173]** Siberia (the Russian Far East): traditionally spoken along the Tatar Strait around the mouth region of the Amur River in Khabarovsk Region and on both the northwestern and the eastern coast of Sakhalin in Sakhalin Province in the Russian Federation. Nivkh consists of Amur Nivkh and Sakhalin Nivkh, of which Sakhalin Nivkh comprises three dialects, geographically identified as North Sakhalin, East Sakhalin and South Sakhalin; North Sakhalin shows features transitional towards Amur Nivkh. In the 1960s and 1970s, the Amur Nivkhs were forced to move to various settlements, and only in the single village of Aleyevka on the Amur River in Komsomol'sk County, a distinct Nivkh community may be said to exist; other villages with Nivkh population include Vlas'yevvo, Baydukovo, Makarovka and Tneyvakh in the same county. On Sakhalin, the original Nivkh villages were liquidated around the same time, and the people were sent to live in three larger settlements: those from the Schmidt peninsula in the north to the village of Nekrasovka and smaller neighbouring villages such as Moskal'vo, Chir-Unvd and Viakhtu in Okha County, those living in the northwestern corner of the island to the village of Rybnovsk (Rybnoye) in the same county, and those inhabiting the eastern shore and the basin of the river Tym' to the town of Nogliki in Nogliki County; a small group lives in the town of Poronaysk in Poronaysk County. In 1989, 1,200 speakers were reported, but the number of active users is probably in the range of a few hundred. There are no younger speakers in the Amur region, and very few, if any, on Sakhalin, though Sakhalin Nivkh appears to be somewhat more vigorous than Amur Nivkh; nevertheless, most speakers are middle-aged or older. Several of the Nivkh speakers who used to live in the territory of modern Poronaysk County in the southern part of Sakhalin were evacuated to Hokkaido, Japan, after 1945; a few of these emigrants survived until recently. There is a written standard based on Amur Nivkh, also used in Nekrasovka, and a recently created literary standard based on the East Sakhalin dialect is being propagated with some success, both in Cyrillic script. Sakhalin Nivkh is severely endangered and Amur Nivkh critically endangered.

**Nogay [143]** Caucasia and Eastern Europe: spoken in several areas in northern Caucasia and southern Russia as well as in the northern parts of the Dobruja region in Romania, and before the deportations to Central Asia after the Second World War around Perekop in the northern steppe area of the Crimea. In 1989, approximately 70,000 speakers were reported from the Russian Federation, where they live in parts of the Karachay-Cherkes Republic, Nogay, Tarumovka, Kizlyar and Babayurt counties in the Republic of Dagestan, Shelkovskaya County in the Republic of Chechenia, and Neftekumsk, Mineral'nyye Vody and Kochubeyevskoye counties in Stavropol' Region. The speakers of Crimean Nogay are counted among the Crimean Tatar. In the Dobruja region in Romania there are up to 25,000 'Tatar' speakers, which covers speakers of both Crimean Tatar and Nogay. Children living in more remote areas are likely to learn Nogay, but many of them shift to other languages at school age. Because of the small, scattered communities, and the overwhelming dominance of

other languages everywhere in its traditional area, Nogay is increasingly endangered. – **Alabugat Tatar [144]** Eastern Europe: south-eastern Russia. Spoken in the village of Severnyy at the station of Ulan-Khol in Lagan' County in the Kalmyk Republic in the Russian Federation. Alabugat Tatar is an outlying dialect of Nogay with deep influences from Tatar and other Turkic languages and, in particular, Kalmyk. There were over 400 members of the ethnic group, presumably all speakers, in 1987. They have traditionally employed the Tatar literary language, and the school operated in Tatar until 1960 when it was shifted to Russian. Because of the small size of the community and the increasing influence of Russian the language must be regarded as definitely endangered. – There are two outlying dialects of Nogay spoken in Astrakhan Province in the Russian Federation, collectively known as **Astrakhan Nogay**; the schools in their areas operated in Tatar until 1962, but since then the Russian influence has been growing and they are now clearly endangered: **Karagash** has been influenced especially by Kazakh. It is spoken in Krasnoyarskiy and Kharabali counties by a few thousand speakers. – **Yurt Tatar** has been influenced especially by Tatar. It is spoken in Volga (Privolzhskiy), Nariman and Volodarskiy counties as well as in the suburbs of the city of Astrakhan by several thousand speakers; a small group known as Kundor Tatars lives in the village of Tuluganovka in Volodarskiy County.

**Norman [70]** France: spoken in the regions of Upper Normandy (Haute-Normandie) and Lower Normandy (Basse-Normandie) by a small portion of the population. Very few if any children learn the language. Even the language of the most competent speakers is likely to be heavily influenced by French. Severely endangered. – There are three outlying dialects of Norman spoken on the Channel Islands, collectively known as **Channel Islands French: Alderney French [auregnais] [71]** was spoken on the island of Alderney in the Bailiwick of Guernsey. The last native speakers died around 1960. Recently extinct. – **Guernsey French [guernesiais] [72]** is spoken on the island of Guernsey in the Bailiwick of Guernsey by perhaps 5,000 speakers, practically all of them middle-aged or older. Severely endangered. – **Jersey French [jèrriais] [73]** is spoken on the island of Jersey, which constitutes the Bailiwick of Jersey, and on the island of Sark in the Bailiwick of Guernsey [*sercquiais*]. In 1989, less than 6,000 speakers were reported from Jersey; the number of speakers on Sark is below 100. There are some younger speakers but most are middle-aged or older. Severely endangered.

**North Frisian [47]** Germany: western coast of Schleswig north of Husum including the Halligen Islands and the adjacent islands of Föhr, Amrum, Sylt, and Helgoland; formerly extended to adjacent Denmark. There are approximately 8,000 speakers; language use is most active on western Föhr and in the village of Risum-Lindholm on the mainland. A small number of children learn the language, but few of them continue to use it in adulthood. Most idiolects are heavily influenced by Low Saxon and German. Severely endangered in Germany; extinct in Denmark.

**North Saami [5]** Northern Europe: Lapland. Spoken in most parts of Finnmark and Troms provinces and in the northernmost corner of Nordland Province in Norway, in Kiruna and Pajala counties and parts of Gällivare and other adjacent counties in Norrbotten Province (Torne Lappmark) as well as, because of resettlements in the twentieth century, also in more southerly regions in Sweden, and in Utsjoki and Enontekiö counties, western parts of Inari County, and Vuotso region of Sodankylä

County in Lapland Province in Finland. Formerly extended to Petsamo region in Murmansk Province in the Russian Federation. Includes three dialect groups, Torne Saami, Finnmark Saami and Sea Saami. In Norway, the number of speakers is above 10,000, in Sweden perhaps 5,000, and in Finland approximately 1,500. In the core area in central Finnmark Province, where North Saami has an official status, many children learn the language and are likely to continue to use the language in adulthood, and in the adjacent parts of Sweden and Finland, there are also child speakers. In other regions, however, North Saami is still giving away to dominant languages. Despite the recent positive trends in parts of its range, North Saami must be regarded as definitely endangered in Sweden, Norway and Finland; extinct in the Russian Federation.

**Northern Altay [148]** Siberia: in the northeastern river valleys of the Altay area in southern Siberia. Northern Altay is a conglomeration of three closely related dialects: **Tuba**, **Kumandy** and **Chalkan**. Tuba is spoken in Turochak, Choya and Mayma counties in the Republic of Altay in the Russian Federation by a couple of thousand people, including few children. Kumandy is spoken in Krasnogorskoye, Solton, Kytmanovo and Tselinnoye counties in Altay Region and in adjacent counties in the Republic of Altay as well as in the towns of Tashtagol and Sheregesh in Kemerovo Province by several thousand people, but only rarely among younger generations. Chalkan (Kuu) is spoken in the villages of Kurmach-Baygol, Suranash, Malyy Chibechen' and Itkuch in Turochak County by approximately 2,000 people. Altogether, there are probably less than 10,000 speakers. The transmission of the language to children is becoming rare but has probably not stopped completely. Southern Altay is used as the written standard. Definitely endangered.

**Northern Khanty [28]** Siberia: spoken in the lower Ob' basin and on its tributaries, mainly in Beloyarskiy and Oktyabr'skoye counties and in the adjacent parts of Berezovo County in Khanty-Mansi Autonomous District and in Shuryuskary County and the southern parts of Ural (Priural'skiy) County in Yamal Nenets Autonomous District in the Russian Federation. Northern Khanty dialects include Ob' Khanty, Shuryuskary and Berezovo Khanty, Kazym Khanty, Sherkaly Khanty and Niz'yam Khanty. In 1989, 14,000 speakers were reported for Khanty as a whole, and since a clear minority of them speak Eastern Khanty, the number of Northern Khanty speakers is possibly more than 10,000. Northern Khanty is still used as a family language, but the boarding school system often forces a shift to Russian among younger generations. Separate written norms in Cyrillic script exist for some of the dialects, but they have very limited use. Definitely endangered.

**Northern Mansi [24]** Siberia: on the western tributaries to the lower Ob', mainly along the Sos'va, and in the central and northern Ural mountains, within the forest zone, mainly in Berezovo County and the northern parts of Sovetskiy County in Khanty-Mansi Autonomous District, formerly also in adjacent parts of the Komi Republic and Yekaterinburg Province in the Russian Federation. Northern Mansi dialects include Sygva Mansi, Sos'va Mansi, Ob' Mansi, and Upper Loz'va Mansi. In 1989, 3,400 speakers were reported for Mansi as a whole; since this figure includes a maximum couple of hundred Eastern Mansi speakers, there are approximately 3,000 speakers of Northern Mansi. There are very few, if any, child speakers; most speakers

are middle-aged or older. There is a written standard in Cyrillic script, but its use is limited to school textbooks. Severely endangered.

**Northern Selkup** (also called **Taz Selkup**) [36] Siberia: in the basin of the river Taz (flowing to the Arctic Ocean), as well as in the Baykha-Turukhan river system, to the west of the upper Yenisey, partly within the tundra zone, mainly in Krasnosel'kup County in Yamal Nenets Autonomous District, the northern parts of Turukhansk County in Krasnoyarsk Region and the northeastern corner of Nizhnevartovsk County in Khanty-Mansi Autonomous District in the Russian Federation. In 1989, 1,800 speakers were reported for all three Selkup languages; the number of Northern Selkup speakers lies somewhere between 1,000 and 1,500, and includes, in the remotest areas, some children. Most children are monolingual speakers of Russian, but Selkup is probably used in some families for internal communication. A written standard in Cyrillic script has recently been reintroduced and is being used, with modest success, in elementary-level school instruction. Severely endangered.

**Olonetsian** [in Olonetsian *livvi*, in Russian *livvikovskij jazyk*, in Finnish *aunus*] [14a] Northern Europe: northwestern Russia. Spoken northeast of Lake Ladoga in the southwest of the Republic of Karelia in the Russian Federation in an area which includes Olonets County (except the eastern corner), the western parts of Pryazha County and the southwestern part of Kondopoga County. In Finland, Olonetsian is spoken by people evacuated, in 1940 and 1944, from former Finnish territories north of Lake Ladoga including Salmi and Impilahti counties, currently the eastern parts of Pitkyaranta (Pitkäranta) County and the southern corner of Suoyarvi (Suojärvi) County in the Republic of Karelia. In the Russian Federation, there are possibly approximately 25,000 speakers. A few children learn the language, but most if not all of them become more fluent in Russian and largely stop using Olonetsian later in life. A literary language written in the Latin alphabet has recently been developed. In Finland, there are a few thousand speakers, all elderly, and even those who are fluent in Olonetsian use mostly Finnish. Definitely endangered in the Russian Federation; critically endangered in Finland.

**Ongkor Solon.** See **Solon**

**Oroch** [in Oroch *naani*] [169] Siberia (the Russian Far East): spoken in the northern section of the Sikhote Alin mountain range, to the east of the lower Amur, now in two isolated groups in Khabarovsk Region in the Russian Federation, one mainly in Vanino County, near the town of Sovetskaya Gavan' along the rivers emptying into the Tatar Strait, including the villages of Datta and Us'ka-Orochskaya on the Tumnin River, and the other on the Amur River in Komsomol'sk County. In 1989, 150 speakers were reported, all middle-aged or older; among the languages of the so-called 'Peoples of the Far North', Oroch has the lowest proportion of native speakers, approximately 20 per cent, in relation to the corresponding ethnic group. Critically endangered.

**Orok** [in Orok *ulta* or *uilta*; in Japanese *Orokko* or *Uiruta*] [172] Siberia: spoken in the northern and central parts of the eastern shore of Sakhalin, in Nogliki, Poronaysk and Aleksandrovsk counties in Sakhalin Province in the Russian Federation, now mainly

in the village of Val in Nogliki County, but many Oroks have moved to the county centres Nogliki and Poronaysk. There used to be Orok speakers also in the southern part of the island, from where people were evacuated to Hokkaido, Japan, after 1945; the descendants of this small emigrant population have by now lost the Orok language. In 1989, eighty speakers were reported, but the actual number is below forty, all middle-aged or older. Among the languages of the so-called ‘Peoples of the Far North’, Orok has the smallest number of native speakers out of the smallest size of the corresponding ethnic group. A project aiming at creating a literary norm for Orok in Cyrillic script and teaching the language at elementary schools has recently been launched in cooperation with Japanese scholars; earlier, the Japanese syllabic script (katakana) has also been used to transcribe Orok material. Critically endangered.

**Oyrat [161]** Mongolia and Central Asia: spoken mainly in Kobdo and Ubsu leagues in the west of Mongolia and in Jungaria region (Bortala, Hoboksar, Tarbagatai and Bayangol) in the north of Xinjiang Uygur Autonomous Region in China, and by smaller groups in the Kukunor region in Qinghai in China, Alashan League in the west of Inner Mongolia in China and the Issyk-Kul’ region in Kyrgyzstan. Oyrat includes nine dialect groups known by tribal names, including Torgut, Khoshut, Ölot and Dörbet. The number of Oyrat speakers can only be estimated because they are officially counted as Mongolians both in Mongolia and China; in Mongolia there are approximately 150,000, in Sinkiang less than 130,000, and in other areas only a few thousand speakers. In the Issyk-Kul’ (Issyk Köl) region, where the Oyrats are officially called Kalmyks, Oyrat is only spoken by the oldest generation, constituting a few of hundred speakers. In most areas, Oyrat continues to be used as a community language but it is under severe pressure from Mongolian, Uygur and Northern Chinese. As a whole, it is increasingly endangered, and in Kyrgyzstan critically endangered. – **Manchurian Ölot** [in Written Mongolian *Yeke Mingghan*, in Modern Mongolian *Yikh Mianggan*] **[162]** Manchuria: still spoken on the eastern bank of the lower Nonni, in Fuyu County in Heilongjiang Province in China; extinct since the early twentieth century in the Imin region of Hulun Buir League in Inner Mongolia. Manchurian Ölot is an outlying dialect of Oyrat. There are probably less than 1,000 speakers, most of them elderly, all in the Nonni region. There have been attempts to support Manchurian Ölot by teaching Standard and Written Mongolian at a local school. Severely endangered.

**Permyak** (also called **Permyak Komi** or **Permian Komi**; the Russian-based term ‘Komi-Permyak’ better avoided) **[21a]** Eastern Europe: east-central Russia. Spoken in Permyak Komi Autonomous District and the adjacent parts of Kirov (Vyatka) and Perm’ provinces in the Russian Federation. In 1989, 110,000 speakers were reported. Quite a lot of children learn the language, but only in remote rural areas they tend to continue to use the language in adulthood. Definitely endangered. Often considered a single language with Komi. – **Yazva Komi** (also called **Eastern Permyak**) **[21b]** an outlying dialect of Permyak, is spoken in the Yaz’va Valley in Krasnovishera County in northeastern Perm’ Province. There are a couple of hundred speakers at best, possibly all elderly, out of a population of some 4,000. Severely endangered.

**Picard [74]** In France, Picard is spoken in the regions of Picardy (Picardie) and Nord-Pas-de-Calais by a small but notable portion of the population. In Belgium, it is spoken by an estimated 200,000 people in the area extending from Tournai to Mons in

the western part of the province of Hainaut. A few children learn the language, and the future of Picard seems slightly less bleak than that of other Oil languages of France. It is nevertheless severely endangered.

**Piedmontese [78]** Italy: spoken in Piedmont Region except the Novara Province, the western Alpine valleys and southern border areas, as well as in the border area of Savona Province in Liguria. There are a couple of million speakers, but particularly in towns and among the young the use of Piedmontese has diminished drastically. Increasingly endangered.

**Pite Saami** (also called *Arjeplog Saami*) [3] Northern Europe: Lapland. Traditionally spoken in most parts of Arjeplog County in Norrbotten Province (Pite Lappmark) in Sweden. There are less than twenty elderly speakers. Nearly extinct.

**Plautdietsch [48]** an outlying dialect of Low Saxon which derives from the concentration of Mennonites in the area of the Vistula Delta in what is now northern Poland and their subsequent emigration to the Ukraine and further to the Americas. Most speakers in the Ukraine were deported during the Second World War to Central Asia and Siberia. Many Central Asian and Siberian speakers have recently emigrated to Germany or other countries. Plautdietsch still has a couple of hundred thousand speakers, including a number of younger people, mainly in the Americas, where the approximate numbers of speakers are for Argentina 4,000, Belize 6,000, Bolivia 60,000 to 80,000, Brazil 8,000, Costa Rica 2,000, Canada 80,000 to 100,000, Mexico 40,000, Paraguay 40,000, Uruguay 2,000, and the United States 10,000. In Europe, Central Asia and Siberia, the transmission of the language to children has all but stopped. Plautdietsch may be regarded as the last remnant of West Prussian Low Saxon. Definitely endangered.

**Poitevin-Saintongeais [68]** France: spoken in the region of Poitou-Charentes and in the department of Vendée in Pays de la Loire region as well as in northern Gironde in Aquitaine region by a small portion of the population. Very few if any children learn the language. Even the language of the most competent speakers is likely to be heavily influenced by French. Severely endangered.

**Pontic Greek [98]** an outlying dialect of Greek which at the beginning of the twentieth century was spoken along the entire Black Sea coast of Turkey and adjacent Georgia: western Pontic was spoken from Inebolu to Ünye, and eastern Pontic in Trebizond (Trabzon, Trapezunt) area and in Chaldia, especially around Gümüşhane. Most speakers in Turkey emigrated to Greece after the First World War; Muslims remained, and are to some extent still found around the towns of Of and Sürmene. There was significant migration to southern Russia and Caucasia in the eighteenth and nineteenth centuries, where Pontic Greek is now healthiest, but many of these populations are now migrating to Greece. In 1989, 40,000 speakers of Greek were reported from the Russian Federation, including 15,000 in Krasnodar Region and the same number in Stavropol' Region; 60,000 speakers were reported from Georgia, and 2,500 from Armenia; the Greek dialects spoken in the regions of Tsalka in Georgia and Alaverdi in Armenia derive partly from Cappadocian Greek but seem to have assimilated to Pontic. The total number of Pontic Greek speakers can only be guessed but it may still

be closer to 200,000 than 100,000. There are probably a few child speakers, but most speakers are middle-aged or older. In Greece, Pontic Greek retains mainly symbolic functions. Definitely endangered in the Russian Federation and Caucasia, severely endangered in Turkey, and critically endangered in Greece.

**Provençal [65]** France: spoken in the departments of Alpes-Maritimes (except the eastern corner), Bouches-du-Rhône, Var and Vaucluse, in the southern parts of Alpes de Haute-Provence, and the eastern parts of Gard, by approximately 200,000 speakers, most of them middle-aged or elderly. A small number of children learn the language, but they usually shift entirely to French at school age. Definitely endangered. Often regarded as a single language with Alpine Provençal.

**Resian Slovene [42]** Italy: an outlying dialect of Slovene spoken in the five villages that constitute the municipality of Resia in the northeastern part of Udine Province in the Autonomous Region Friuli-Venezia Giulia by the majority of the approximately 1,400 inhabitants of the municipality. In the westernmost villages of San Giorgio, Gniva and Prato there are few, if any child speakers, but in the villages of Oseacco and Stolvizza children are still learning the language to some extent. The language of younger men in particular is quite heavily influenced by Italian. Definitely endangered.

**Romani [104]** dispersed in many European countries, most densely in East-Central and Eastern Europe and in the Balkans. Many speakers were murdered by Germans during the Second World War. Seven Romani dialects are still spoken: (1) *Vlach* in Albania, Bosnia and Hercegovina, Hungary, the Republic of Moldova, Romania, the Ukraine, Yugoslavia and neighbouring countries; (2) *Balkan Romani* in Bulgaria, Greece, the Republic of Moldova, the Republic of Macedonia, Romania, Turkey, the Ukraine and Yugoslavia; (3) *Welsh Romani* in Wales; (4) *Finnish Romani* in Finland; (5) *Sinte* in Austria, Croatia, the Czech Republic, France, Germany, Italy, the Netherlands, Poland, Slovenia, Switzerland and Yugoslavia; (6) *Carpathian Romani* in the Czech Republic, Hungary, Poland, Romania, Slovakia, and the Ukraine; (7) *Baltic Romani* in Belarus, Estonia, Latvia, Lithuania, Poland, Russia and the Ukraine. There are at least three and a half million speakers in total (see Matras 2002). In 1989, 130,000 speakers were reported from the Russian Federation, 30,000 from the Ukraine, 10,000 from the Republic of Moldova, and 30,000 from other parts of the Soviet Union. Welsh and Finnish Romani have few speakers left; the situation of Sinte may actually be similar; other dialects have large numbers. There are many children who learn Romani, and they would have all chances to continue to use it through adulthood, were it not that practically all national governments are hostile to the Romani language and culture. In many countries, such discriminatory policies have already led to a situation where children no longer learn the language. Increasingly endangered.

**Romansh [in German often *Bündnerromanisch*] [126]** Switzerland: Graubünden (Grischun), mainly in Surselva, Seumeir, and Unterengadin. Perhaps little more than half of the official number of 65,000 speakers (1986) use the language actively. Many children learn the language, but most of them become more fluent in German and may not use Romansh actively later in life. Increasingly endangered.

**Rutul** (also known as *Mukhad*) [126] Caucasia: southern Daghestan and northern Azerbaijan. Spoken in seventeen villages, including Rutul, Shinaz, Ikhrek, Myukhrek, Luchek, Amsar and Borch, in Rutul County and in the village of Khnov in Akhty County in the Republic of Daghestan in the Russian Federation, and in adjacent villages in Kakh (Qax) and Sheki (Šäki) counties in Azerbaijan. In 1989, 19,000 speakers, almost all in Daghestan, were reported from the Soviet Union, including many children, but because of the small size of the community and the dominant position of Russian, Azerbaijani and Lezgian, the language must be regarded as definitely endangered. Until recently, there was no literacy in it, but a Rutul primer was published a few years ago.

*Sassarese Sardinian*. See *Corsican*

**Saterlandic** (also called *Sater Frisian* or, earlier more correctly but now possibly creating confusion with neighbouring dialects of Low Saxon, *East Frisian*) [in Saterlandic. *seeltersk*] [46] Germany: spoken in the towns of Strücklingen (Strukelje), Ramsloh (Romelse) and Scharrel (Schäddel) in Saterland region west of Oldenburg; East Frisian was also spoken on the island of Wangerooge till the early twentieth century. Although the commonly cited figure is 2,000, there are probably much less than 1,000 speakers, including few if any child speakers. While there may be some younger speakers, most are middle-aged or older. All idiolects are likely to be heavily influenced by Low Saxon and German. Severely endangered.

**Scottish Gaelic** [55] Scotland: rural areas of the Western Isles (Lewis, Harris, North Uist, South Uist, Barra) and Skye, and a few locations in the rest of the Inner Isles and the Highland mainland (mainly Sutherland, Ross and Cromarty, Inverness, and Argyll counties). There are 20,000 to 30,000 active users; more than 50,000 others claim knowledge of the language. Scottish Gaelic is also spoken elsewhere in Scotland, and by one to several thousand speakers in immigrant communities in Nova Scotia and Prince Edward Island in Canada. A number of children learn the language, but there are serious problems in language maintenance even in the core areas. Definitely endangered.

**Shor** [149] Siberia: spoken along the rivers flowing from the southwestern slopes of the Kuznetskiy Alatau mountains in the southern parts of Kemerovo Province in the Russian Federation. Shor is a conglomeration of two historically distinct dialects, identified by river names as Kondoma Shor (the southern dialect, close to Northern Altay) and Mrassu Shor (the northern dialect, close to Khakas). The Shor-speaking territory, known as Shoria, briefly held the status of an autonomous district (1925–39), but was subsequently turned into one of the greatest industrial regions of the Soviet Union (the so-called Kuzbass region). In 1989, less than 10,000 speakers were reported. There are probably no child speakers, except possibly in a few rural families, and most speakers are middle-aged or older. The Shor community is surrounded by Russian-speaking immigrants, but there have recently been signs of increased national and linguistic consciousness. A written standard in Cyrillic script is based on Kondoma dialect. Severely endangered.

*Sinkiang Dagur*. See *Dagur*

**Skolt Saami** [in Finnish *kolтта*] [7] Northern Europe: Lapland and Kola peninsula. Spoken in Sevettijärvi region in Inari County in Lapland Province in Finland, mainly by people evacuated from former Finnish territory of Petsamo, now Pechenga County in Murmansk Province in the Russian Federation. Earlier spoken also in the western parts of Kola County in western Murmansk Province, from where the speakers were forcibly translocated to Lovozero, the centre of Lovozero County. The number of speakers in Finland is around 300. There were approximately twenty speakers in the Russian Federation in the early 1990s. Formerly also spoken in easternmost Finnmark Province of Norway. In Finland, very few children learn the language, and in the Russian Federation, there are only elderly speakers. Skolt Saami has been used as a literary language since the 1970s. It is severely endangered in Finland, nearly extinct in the Russian Federation, and extinct in Norway.

**Solon** [in Chinese pinyin *Suolun* or *Suolun Ewenke*] [166] Manchuria: spoken in central and northwestern Manchuria in China, in Morin Dawa Daur and Ewenki autonomous banners and adjacent areas of neighbouring banners in Hulun Buir League in northeastern Inner Mongolia (in the Imin basin) as well as in Heilongjiang Province, mainly in Nahe (Nehe) County (in the Nonni basin), west and south of the Evenki. Solon speakers are included in the official Ewenke nationality; cf. Evenki. There are perhaps fewer than 10,000 speakers, out of 21,000 ethnic group members, and in many places they are over fifty years old. As to child speakers, the situation varies greatly, with some groups in Heilongjiang adopting other languages (Dagur, Chinese). On the Inner Mongolian side the language is generally being retained better, with some children learning it, depending on locality. Most if not all speakers also speak Dagur, and many speak Northern Chinese. The function of literary languages is fulfilled by both Chinese and Modern Written Mongolian; attempts have also been made to write Solon in various systems of transcription (Mongolian, Roman, IPA). Severely endangered. – **Ongkor Solon** (also called *Sinkiang Solon* or *Turkestan Solon*) [167] Sinkiang: spoken in the Ili (Yili) region, in Kazakh Autonomous District in Xinjiang Uygur Autonomous Region in China. Ongkor Solon is an outlying dialect of Solon: Solon speakers were translocated to Sinkiang from Manchuria together with Dagurs in 1763. The last fluent speaker, of the male sex, was seventy-nine in 1990, and his death was announced recently. There may still be a few people who know some isolated phrases or words. Recently extinct.

**Sorbian** (also called *Lusatian* or *Wendish*) [44] Germany: Lusatia (Lausitz [Łužica or Łužyca]) region around Cottbus (Chośebuz) and in the Spreewald (Błota) area in Brandenburg (lower Lusatia) and in and around Bautzen (Budyšin) in Saxony (upper Lusatia). There are two literary languages, Lower Sorbian and Upper Sorbian, based on different dialects but the language area is unified by a large area of transitional dialects, which speaks for the recognition of a single spoken language. Although the number of Sorbian speakers is occasionally reported as high as 70,000 or even 110,000, the factual number is much lower, and lies somewhere between 20,000 and 30,000; about one third in lower and two thirds in upper Lusatia. Among the Upper Sorbian Catholics, constituting some 10,000 people between the towns of Bautzen, Hoyerswerda and Kamenz, most children still learn the language and continue to use it in adulthood unless they are forced to move away because of unemployment, while elsewhere shift to German has been proceeding rapidly since the Second World War, so that there are

now very few child speakers among the Upper Sorbian Protestants, and none among the Lower Sorbians who are all Protestants. Most active speakers among Protestants are therefore elderly, although Sorbian is still studied as a second language at school starting from the age of fifteen. Contrary to recent trends among European minority languages, Sorbian has lost much of its former public support after the unification of Germany. Sorbian schools are being closed one after another, officially for financial reasons, but anti-minority feelings are apparent in the background. For instance, the government of Saxony is currently planning to close down the Sorbian Secondary School at Crostwitz (Chrósćicy), which would severely disrupt, if not destroy, the Upper Sorbian school network. Definitely endangered; in lower Lusatia, severely endangered.

**South Saami [1]** Northern Europe: Lapland and central Scandinavia. Spoken in Vilhelmina County and parts of neighbouring counties in Västerbotten Province (Åsele Lappmark), in Strömsund, Krokom, Åre, Berg and Härjedalen counties in Jämtland Province, and in Idre region in Älvdalen County in Dalarna Province in Sweden, and in southern parts (up till Rana) in Nordland Province, many parts of Nord-Trøndelag and Sør-Trøndelag provinces, and in Engerdal region in Hedmark Province in Norway. There are a few hundred speakers, many of whom prefer Swedish or Norwegian, out of a much larger ethnic population. In some families, children may learn the language, but it is nevertheless severely endangered.

**Southern Altay [147]** Siberia: in the central and southwestern river valleys of the Altay area in southern Siberia. Southern Altay is a conglomeration of three closely related dialects: *Altay*, *Telengit*, and *Teleut* (or *Telengut*) [in Chinese pinyin *Tielingute*]. In 1989, a combined figure of 61,000 speakers for Southern Altay and Northern Altay was reported, which means that Southern Altay is spoken by approximately 50,000 people, as long as the figure is not notably inflated. Southern Altay is mainly spoken in the Republic of Altay in the Russian Federation. Teleut specifically is still spoken in Shebalino County in the Republic of Altay, but the bulk of Teleut speakers were dispersed by political developments in the seventeenth and eighteenth centuries, and Teleut is now mainly spoken to the north of the actual Altay region, concentrated in the Bachat region in Belovo, Gur'yevsk and Novokuznetsk counties in the western part of Kemerovo Province and in the Chumysh region in Tselinnoye County in the eastern part of Altay Region. A small diaspora group (probably since the eighteenth century) of Teleut origin has been registered within the Altai District of northern Sinkiang, China, but it is not known if they continue to speak the language. Southern Altay is probably still being learnt by some proportion of children, and even in the case of the widely dispersed Teleut communities consisting of a couple of thousand people, the transmission of the language to children has probably not stopped entirely. Triggered by both social and ecological problems, there is currently a rise of nationalism, which may improve the position of the native language. The speakers of Teleut, in particular, have recently shown interest in national renaissance, which might lead to the revigoration of the native language. Increasingly endangered.

**Southern Khanty [29]** Siberia: formerly spoken in the lower Irtysh basin and on its tributaries, within Uvat County in Tyumen' Province and Konda and Khanty-Mansiysk counties in Khanty-Mansi Autonomous District in the Russian Federation. Dialects included Dem'yanka Khanty, Konda Khanty and Irtysh Khanty. There are

probably no speakers left, but no confirmation is available; in any case, the transmission of the language to children stopped long ago. Southern Khanty was assimilated to both Siberian Tatar and Russian. Possibly extinct.

**Southern Mansi** (also called *Tavda Mansi*) [27] Siberia: last speakers lived in the region of the lower Tavda, a tributary to the lower Tobol'-Irtys, in the border area of Nizhnyaya Tavda County in Tyumen' Province and Tavda County in Yekaterinburg Province in the Russian Federation. Earlier Southern Mansi was spoken over a much larger area covering the southern parts of Yekaterinburg Province and adjacent parts of neighbouring regions. Southern Mansi became extinct before the middle of the twentieth century. The language, as recorded from the last generations of speakers, reveals interference from Siberian Tatar, but the very last speakers seem to have adopted Russian.

**Southern Selkup** [38] Siberia: spoken mainly in the basin of the river Ket', an eastern tributary to the upper Ob', in Kolpashevo and Upper Ket' (Verkhneketskiy) counties in Tomsk Province and adjacent parts of Krasnoyarsk Region in the Russian Federation; formerly spoken further to the south. Consists of two dialect groups, Upper Ob' Selkup, which is practically extinct, and Ket' Selkup, which still has a small number of speakers, probably much less than one hundred, all elderly. Critically endangered.

**Svan** [106] Caucasia: northwestern Georgia. Spoken in an area known as Svanetia, mainly within Mestiya and Lentekhi counties in Georgia, and possibly in small enclaves in the adjacent regions in the Kabard-Balkar Republic in the Russian Federation. The number of speakers is usually given as 40,000, but in reality it may be down to a few thousand, because many people who have shifted to the closely related Georgian still claim Svan ethnicity. While the situation is poorly known, it is possible that there are very few child speakers left. There are certainly some younger speakers, but most speakers appear to be middle-aged or older. There is no literacy in Svan, but Georgian is used as a literary language. Definitely endangered.

**Tat** [102] Caucasia: northern Azerbaijan. Spoken in Syazan, Divichinsk, Kuba, Konakhkend, Semakh and Ismail regions in the northeast of Azerbaijan as well as on Apsheron peninsula and Baku. Apart from the Jewish Tats traditionally speaking the closely related but distinct Juhur (Judeo-Tat) language, the Tat nationality is divided into two groups, the Muslim Tats and the Christian (or Armenian) Tats. Many of the Christian Tats settled in Daghestan already in the eighteenth and nineteenth centuries, and the people in the last Christian Tat communities in Azerbaijan, in the villages Kilvar in Divichinsk region and Madras in Shemakh region, were forced to move to Armenia in the late 1980s. The 1989 census reports 5,000 speakers of Tat in Azerbaijan, 4,000 of them in Baku, but these figures are far from trustworthy. There may possibly still be some child speakers, but a shift to Azerbaijani has proceeded rapidly among Muslim Tats. Tat is not written, but the speakers in Azerbaijan use Azerbaijani in school and for wider communication. Definitely endangered.

**Ter Saami** [10] Northern Europe: Kola peninsula. Earlier spoken in the eastern parts of Lovozero County in Murmansk Province in the Russian Federation, from where the speakers were forcibly translocated to Lovozero, which lies outside the native

territory. There were six speakers in the early 1990s, all elderly. Some descendants of Ter Saami speak Kildin Saami, but most have shifted to Russian. Nearly extinct.

**Tindi [117]** Caucasia: western Daghestan. Spoken in the villages of Tindi, Angida, Aknada, Echeda and Tissi in Tsumada County in the Republic of Daghestan in the Russian Federation. The number of speakers lies somewhere between 4,000 and 6,000, with relatively few child speakers. Because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. There is no literacy in it.

**Tofa [155]** Siberia: originally spoken by nomadic groups on the northern slopes of the Eastern Sayan mountains, now in the villages of Alygdzher, Nerkha and Verkhnyaya Gutara in Nizhneudinsk County in Irkutsk Province in the Russian Federation, to the south of the city of Nizhneudinsk. In 1989, 300 speakers were reported, and the actual figure was probably lower, and currently there are perhaps no more than forty fluent speakers plus the same number of passive speakers in Alygdzher and Nerkha. All known fluent speakers are above forty years of age, but the situation may be slightly better in Verkhnyaya Gutara. Critically endangered.

**Töitschu [50]** Italy: an outlying dialect of Alemannic or, more precisely, of Walser (Highest Alemannic), spoken in the village of Issime (Eischeme) in the upper Lys valley. Another Walser dialect, Titsch, is spoken in the villages of Gressoney-Saint-Jean and Gressoney-La-Trinité, but it appears to be very closely related to and in constant contact with cross-border Alemannic in Switzerland; the same may be true of other varieties of Walser spoken outside Valais (Wallis) Canton, notably in the village of Bosco-Gurin in Ticino Canton as well as in Graubünden and parts of Austria. Töitschu is a typical, rather independent outlying dialect that has been influenced by Francoprovençal and Piedmontese. The combined number of speakers of both Töitschu and Titsch in the Aosta Valley is reportedly 600. Definitely endangered.

**Trukhmen** (also called *Caucasian Turkmen*) [137] northern Caucasia and Eastern Europe: an outlying dialect of Turkmen spoken since the eighteenth century in southern Russia, mainly in Stavropol' Region and Astrakhan Province in the Russian Federation. A figure of 18,000 speakers has been reported, and a few children probably learn the language, but the situation is poorly known. Definitely endangered.

**Tsakhur [127]** Caucasia: northern Azerbaijan and southern Daghestan. Spoken in sixteen villages in Zakatala (Zaqatala) and Kakh (Qax) counties in Azerbaijan as well as in thirteen villages in Rutul County in the Republic of Daghestan in the Russian Federation. In 1989, 13,000 speakers were reported from Azerbaijan and 6,000 from the Russian Federation, but the actual figures may be even higher. Most children are speakers, but because of the relatively small size of the community and the dominant position of Azerbaijani and Russian, the language must be regarded as definitely endangered. Tsakhur was used as a literary language in the mid-1930s and again in the 1990s, and in recent years it has been studied as a subject in local schools.

**Tsakonian [101]** Greece, eastern Peloponnese: southern Tsakonian is still spoken in villages around Leonidio, but practically all speakers are elderly, while northern

Tsakonian barely survives in the villages of Kastanitsa and Sitena. There was also an outlying dialect called Propontis Tsakonian spoken by colonists in northeastern Turkey, in the villages of Havoutsi and Vatka on the mouth of the Gönen river, but after the speakers were settled in Greek Macedonia following the 1922 population exchanges, it became extinct. There are a couple of hundred people with some fluency in Tsakonian at best, mainly in Leonidio region but still a few in Kastanitsa as well. Northern Tsakonian has undergone earlier influence from Greek than southern Tsakonian, but there is also significant influence discernible in southern Tsakonian. Critically endangered.

***Tsez* (or *Dido*) [119]** Caucasus: western Daghestan. Spoken in several villages, including Khutrakh, Kidero, Khupri, Shaitl, Mikok, Tsebari, Asakh, Shapikh and Sagada, in Tsunta County in the Republic of Daghestan in the Russian Federation, and in a number of expatriate communities elsewhere in Daghestan as well as in Turkey. The number of speakers is probably between 7,000 and 8,000. Because of the small size of the community and the dominant position of Avar, the language must be regarded as definitely endangered. Until recently, there was no literacy in it, but a Tsez primer was published in 1993.

***Tundra Enets*** (also called *Madu* or *Somatu*) [32] Siberia: in the tundra zone on the lower Yenisey, now concentrated in the village of Vorontsovo in Ust'-Yeniseysk County in Taymyr (Dolgan and Nenets) Autonomous District in the Russian Federation. Historically, there has occurred a northward movement comparable to that of Forest Enets. The number of speakers is perhaps around thirty, and all of them are middle-aged or older. The last speakers are bilingual in Russian, but there may also be individuals with a knowledge of Nganasan. Critically endangered.

***Tundra Nenets*** [34] Siberia and northern Russia: spoken in northwestern Siberia across a vast area in the Russian Federation, covering Yamal, Nadym and Taz counties and the northern parts of Ural (Priural'skiy) County in Yamal Nenets Autonomous District as well as most parts of Ust'-Yeniseysk County in Taymyr (Dolgan and Nenets) Autonomous District, and in the northernmost parts of European Russia including Nenets Autonomous District, the Kolguyev Island, the northern parts of Mezen' County and formerly the Novaya Zemlya Islands in Arkhangel'sk Province, as well as Vorkuta Territory and at least formerly the northern parts of Inta Territory and Usinsk, Ust'-Tsil'ma and Izhma counties in the Komi Republic. In 1989, 27,000 speakers were reported for the two Nenets languages; of these approximately 25,000 were speakers of Tundra Nenets. In Siberia, many children learn the language, but often shift to Russian at school age, yet most young people are still fluent in the language. On the European side, very few children learn the language, young people tend to prefer Russian, and most speakers there are middle-aged or older. Definitely endangered.

***Tundra Yukagir*** (also called *Northern Yukagir* or *Wadul*) [176] Siberia: spoken in the tundra zone, in a belt extending from the lower Indigirka in the west close to the lower Kolyma basin in the east, in the Republic of Sakha (Yakutia) in the Russian Federation, now largely concentrated in the villages of Andryushkino and Kolymskoye in Lower Kolyma (Nizhnekolym'skiy) County. Previously Tundra Yukagir was spoken in

a much wider area in the Lena-Yana-Indigirka-Kolyma region. In 1989, 400 speakers were reported for both Yukagir languages, but the figure is clearly inflated. There are fewer than 150 speakers of Tundra Yukagir, including elderly people who have it as their first language and middle-aged people who are typically more fluent in Russian and Yakut. In Andryushkino there may be a few younger speakers. All the remaining speakers are multilingual in Russian, Yakut, Even or Chukchi. Attempts are currently being made to create a written standard in Cyrillic script, with either Russian or Yakut-based orthographical principles, for both Tundra Yukagir and Forest Yukagir. Critically endangered.

**Ubykh [109]** originally spoken in northern Caucasia, on the Black Sea coast north from Khosta in what is now the southernmost part of Krasnodar Region, but practically the entire Ubykh nation, some 50,000 people, migrated to Turkey in 1864. The last speaker, Tevfik Esenç, from the village of Hacı Osman Köyü near the Sea of Marmara, died in October 1992. Recently extinct.

**Udege** (also spelled *Udihe* or *Udehe*) [168] Siberia (the Russian Far East): in the southern and central sections of the Sikhote Alin mountain range, to the east of the Ussuri river, now mainly concentrated in four villages: Krasnyy Yar in Pozharskoye County and Agzu in Terney County in Maritime (Primor'ye) Region, and Gvasyugi in the County of Lazo and Arsen'yevo (Rassvet) in Nanay County in Khabarovsk Region in the Russian Federation. In 1989, under 500 first-language speakers out of 1,900 ethnic group members were reported, but the actual number is currently much smaller, possibly below 100, all of them elderly. A shift to Russian has proceeded in Udege communities faster than almost anywhere else in Siberia, and many characteristic features of the language have been lost. Earlier there were speakers of Udege between the Ussuri and Sungari basins in northeastern Manchuria in China, locally known as *Kyakala* [in Chinese pinyin *Qiakala*]. Manchurian Udege became extinct probably in the early twentieth century, but no exact date is available; it was initially replaced by Manchu and finally by Northern Chinese. The Udege were earlier also called *Tazy*, a term which today primarily refers to a local Chinese-speaking population. Critically endangered in the Russian Federation, extinct in China.

**Udi [131]** Caucasia: northern Azerbaijan and eastern Georgia. Earlier spoken across a large area in northern Azerbaijan, but now mainly in the villages of Vartashen (Vardashen, currently Oguz) in Vartashen County and Nij (Niç) in Kutkashen (currently Qabala) County, recently also in Sheki (Şäki) County and part of Zakatali County. Since the 1920s Udi has also been spoken in the village of Okt'omber (formerly Zinobiani) in Kvareli County in Georgia, the dialect of which is closely related to that of Vartashen from where the speakers fled. In 1989, 5,500 speakers were reported from Azerbaijan and a few hundred from Georgia, but the actual numbers may be lower; many Udi may have emigrated from Vartashen to Russia. There are relatively few child speakers, and the language is definitely endangered. Literary experiments took place in the nineteenth century, in the 1930s, and again in recent years. Until recently, Azerbaijani, Georgian and Russian were used exclusively in school, and they continue to be used for wider communication.

**Udmurt [21]** Eastern Europe: east-central Russia. Spoken in the Udmurt Republic and parts of Tatarstan, the Republic of Mariy-El, Bashkortostan, and Kirov (Vyatka) and

Perm' provinces in the Russian Federation. In 1989, 520,000 speakers were reported. Many children learn the language, but only those living in remote rural areas continue to use it actively. Increasingly endangered.

**Ulcha** (also spelled **Olcha**) [in Ulcha *naani*] [171] Siberia (the Russian Far East): spoken in the lower Amur basin, in Ul'cha County in Khabarovsk Region in the Russian Federation. In 1989, less than 1,000 first-language speakers out of 3,200 ethnic group members were reported, all middle-aged or older, and the shift to Russian continues to proceed rapidly. Critically endangered.

**Ume Saami** [2] Northern Europe: Lapland. Traditionally spoken in Arvidsjaur County and the southeastern corner of Arjeplog County in Norrbotten Province (Pite Lappmark) and in Malå and Sorsele counties and the northern Tärna region in Storuman County in Västerbotten Province (Lycksele Lappmark) in Sweden. Formerly also spoken in Rana County in Nordland Province in Norway. There are less than twenty speakers, all elderly. Nearly extinct in Sweden; extinct in Norway.

**Urum** (or *Greek Tatar*) [136] Eastern Europe and Caucasia: an outlying dialect of (Crimean) Turkish originally spoken in the south of the Crimea, now in a few villages in Donets'k (Donetsk) Province in the southeast of the Ukraine and in some places, notably Trialeti, in Georgia and perhaps elsewhere in Caucasia. The number of speakers is not known exactly, but may range in some tens of thousands. A language shift to dominant languages has proceeded rapidly, leaving few child speakers. Definitely endangered.

**Veps** [16] Northern Europe: northwestern Russia. Veps comprises three dialect areas: Northern Veps is spoken in the eastern part of Onega (Prionezhskiy) County in the southeast of the Republic of Karelia in the Russian Federation; Central Veps is spoken in a larger area across the boundary of St Petersburg (Leningrad) and Vologda provinces, mainly in the southern part of Podporozh'ye County, the eastern corner of Tikhvin County and the northeastern part of Boksitogorsk counties in St Petersburg Province, and the western parts of Babayevo and Vytegra counties in Vologda Province; Southern Veps is spoken in the southeastern part of Boksitogorsk County. In 1989, 6,000 speakers were reported, but the figure may actually be slightly too low. In a couple of villages, some children learn the language, but many stop using it at school age. Veps was used as a literary language in the mid-1930s and again in the 1990s, and in recent years it has been studied as a subject in local schools. Severely endangered.

**Vojvodina Rusyn** (or *Bachkan*) [45] the Balkans. Vojvodina Rusyn can be historically seen as an outlying dialect of Eastern Slovak with Rusyn influences, but it is now better regarded as a distinct language. It is spoken by less than 20,000 speakers in the Bačka region in Vojvodina Province in Serbia. There are probably some child speakers but also various negative factors. Definitely endangered.

**Võro-Seto** (or simply *Võro*) [in Estonian *Võru Keel*] Northern Europe: Estonia and northwestern Russia. Spoken in the southeastern corner of Estonia and in Pechory County in the rest of Pskov Province in the Russian Federation by approximately

50,000 speakers of all ages, but younger people are shifting to the majority language. Definitely endangered.

**Vote** [in Russian *vodskij jazyk*, in Estonian *vadja keel*, in Finnish *vatja*] (variants such as ‘Votic’ better avoided) [12] Northern Europe: northwestern Russia. Spoken in a few inland villages south of the Gulf of Finland in Kingisepp County in the west of St Petersburg (Leningrad) Province in the Russian Federation. There are probably less than twenty-five speakers left now, concentrated in the villages of Krakol’e (Jögöperä) and Peski-Luzhitsy (Liivtšülä-Luuditsa), until recently also in Kotly (Kattila) and Mezhniki (Rajo). All of them are elderly, and prefer Russian in daily communication. The above mentioned villages belong to the area of the Western dialect; the last speakers of Eastern Vote, in Icipino (Itšäpäivä), died probably in the 1960s, and those of the distinct, Ingrian-influenced dialect of Kurovitsy (Kukkusi) in the 1980s. All modern idiolects are heavily influenced by Ingrian, Finnish and Russian. Nearly extinct. – An outlying dialect known as *Krevin*, originally spoken by prisoners captured in the fifteenth century, survived in the territory of present-day Latvia until the early nineteenth century.

**Walloon** [75] In Belgium, Walloon is spoken in the greater part of the province of Liège, in the southern part of the province of Brabant, in the province of Namur, in the northern part of the province of Luxemburg and in the eastern part of the province of Hainaut by an estimated 600,000 people, although the number of active speakers may be half that number. In France, it is spoken in the north of the department of Ardennes (town of Givet) by a small number of people, mostly elderly. In Luxembourg, it was formerly spoken in two or three villages (Doncols, Sonlez), where the last speakers died in the 1970s. In Belgium, a number of children learn the language, but many are likely to shift entirely to French. Increasingly endangered in Belgium; severely endangered in France; extinct in Luxemburg.

**Western Mansi** [26] Siberia: formerly spoken in the region of the source rivers of the Tavda, a tributary to the lower Tobol’-Irtys, within Ivdel’ and Gari counties in Yekaterinburg Province in the Russian Federation. Dialects included Pelym Mansi, Middle and Lower Loz’va Mansi and Vagil’sk Mansi. There are probably no speakers left, but when the last ones died is not known. Possibly extinct.

**Western Mari** (often called *Hill Mari* after the largest dialect group) [19] Eastern Europe: east-central Russia. Spoken in westernmost parts of the Republic of Mariy-El and adjacent parts of Kirov (Vyatka) and Nizhniy Novgorod provinces in the Russian Federation. There are probably less than 50,000 speakers, including few children; cf. Eastern Mari; interference from Russian is generally stronger than in Eastern Mari. Definitely endangered.

**Yazva Komi** (Eastern Permyak). See *Permyak*

**Yiddish** (or *Judeo-German*) [49] the traditional language of Ashkenazi Jews, Yiddish was earlier spoken in large areas mainly in Eastern and East-Central Europe, but most speakers were murdered by Germans during the Second World War. Before the war, there were 11–12 million speakers, while in the 1960s, 4 million were estimated. In

Europe, Yiddish is currently spoken in a few places in Belarus and the Ukraine, as well as by a small number of individuals in Alsace, the Netherlands and Switzerland, and in Jewish Autonomous Province (capital Birobijan) in eastern Siberia in the Russian Federation. In 1989, 150,000 speakers were reported from the Soviet Union, but the number of actual users was probably much smaller. Most speakers now live in North America and Israel. More than 1,000,000 people in North America and approximately 200,000 in Israel have knowledge of Yiddish, and there are several communities where children still learn the language, but the general trend is alarming. Definitely endangered in Europe, increasingly endangered worldwide.

*Yug* (also called *Southern Ket* or *Sym Ket*; cf. Ket) [175] Siberia: originally spoken in the basins of the rivers Sym, Kas and Dubches, western tributaries to the middle Yenisey opposite to the Podkamennaya Tunguska, later in the villages of Vorogovo in Turukhansk County and Yartsevo in Yeniseysk County in Krasnoyarsk Region in the Russian Federation. There are probably no speakers left. Possibly extinct.

*Yurt Tatar*. See *Nogay*

## Acknowledgements

I wish to express my gratitude to all colleagues of mine who have provided information for the purpose of this study, first to Juha Janhunen, whose work forms the basis of much of the Siberian and Manchurian data, and further to a large number of specialists, including Jarmo Alatalo (Selkup), Ole Stig Andersen (Ubykh), Lars-Göran Andersson (Dalecarlian), Viktor D. Atknine (Evenki), Peter Bakker (Romani and mixed languages), Mariana Bara (Istro-Romanian, Aromanian and Megleno-Romanian), Sarah Benor (Jewish languages), Martin Böhler (Tsakonian), Jørgen Giorgio Bosoni (Lombard), Wayles Browne (Rusyn, Vojvodina Rusyn), Juris Cibuls (Latgalian), Xavier F. Conde (Galician), Karen H. Ebert (North Frisian), Onno Falkena (West Frisian), Andrei Filchenko (languages of western Siberia), A. N. Garkavec (Urum and other Turkic languages), Luís Gomes (Asturian-Leonese), Godehard Gottwald (Sorbian), Ekaterina Gruzdeva (Nivkh), Reinhard F. Hahn (Low Saxon and Plautdietsch, Sorbian, languages of Sinkiang), Mikko Hakalin (Cypriot Arabic, Karaim, Old Prussian), Panu Hallamaa (Eskimo-Aleut and Finnic), K. David Harrison (Chulym Turk, Dukha, and Tofa), Christian Heinen (Luxemburgian), Eugene Helimski (Enets and Nganasan), Lorint Hendschel (Walloon), Katalin Henriksson (Csángó Hungarian), Davyth Hicks (Cornish, Manx), Stephanie Hughes (Rhenish Franconian), Geraint Jennings (Jersey French), Seth Jerchow (Italkian and other Jewish languages), George Jochnowitz (Jewish languages), Thede Kahl (Megleno-Romanian), Timo Kalmu (languages of eastern Siberia and northern China), Paweł Kaşk (Kashubian), Olga Kazakevitch (Selkup and other Siberian languages), Aleksandr E. Kibrik (Alutor), Tomas Kindahl (Rusyn), Olavi Korhonen (Saami in Scandinavia), Iacovos Koumi (Cypriot Greek), Lars-Gunnar Larsson (Saami in Scandinavia), Larisa Leisiö (Nganasan), Marina Lublinskaya (Tundra Nenets and Nganasan), Walter Lupani (Germanic in Italy), Kaur Mägi (Forest Nenets), Peter Mansfield (Tsakonian), Kazuto Matsumura (languages of the Russian Federation), Lars Narvselius (Scanian), Nick Nicholas (Hellenic languages and languages of Greece), Iliia Nikolaev (Ingrian), Irina Nikolaeva (Udege), Karl Pajusalu (Võro-Seto), Annika Pasanen (Inari Saami and Karelian), Gabriela Popa (Megleno-Romanian), Raija Pyöli (Olonetsian), Elisabetta Ragagnin (Dukha), Leif Rantala (Saami in Kola Peninsula), Torkel Rasmussen (Saami), Gabriele Romagnoli (Molise Croatian), Esa-Jussi Salminen (Udmurt), Gábor Sandi (Francoprovençal), Elena K. Skribnik (Buryat), Sven-Erik Soosaar (Tundra Nenets), Han Steenwijk (Resian Slovene), Jan-Olof Svantesson (Mongolic languages), Jon Todal (Saami), Eva Toulouze (Forest Nenets), Tom Trier (Rusyn), Manuel Trinidad (Extremaduran), Trond Trosterud (languages of Scandinavia), Nikolai Vakhtin (Siberian languages), Philippe Verdy (Gallo), Tiit-Rein Viitso (Livonian, Vote), Dusan Vitanovics (Rusyn), Chris Wilson (Cornish), David Wilson (Scots), Peter Wiens (Plautdietsch), and Birger Winsa (Torne Valley Finnish). I also wish to thank all the other scholars who have helped me with more general problems connected with this study, notably Daniel Abondolo, Florian Blaschke, David Bradley, Lyle Campbell, Paul Fryer, Victor Golla, Riho Grünthal, Joachim Otto Habeck, Michael E. Krauss, Jarel Kuga de Deaton, Ulla-Maija Kulonen, Ago Künnap, Ariadna I. Kuznecova, Jouni Maho, Lubor Mojdl, Nicholas Ostler, Donall Ó Riagain, Merja Salo, Pekka Sammallahti, Tove Skutnabb-Kangas, Seppo Suhonen, and the late Larry Trask and Stephen A. Wurm. Obviously, none of the above mentioned should be held responsible for any errors found in this presentation.

## Bibliography

- Arctic languages: an awakening* (1990) ed. Dirmid R. F. Collis, Paris: Unesco.
- Atlas of the world's languages* (1994) general eds Christopher Moseley and R. E. Asher, London and New York: Routledge.
- Atlas of the world's languages in danger of disappearing*, 2nd edn, revised, enlarged and updated (2001) ed. Stephen A. Wurm, cartographer Ian Heyward, Paris: Unesco.
- Campbell, Lyle (1998) *Historical linguistics: an introduction*, Edinburgh: Edinburgh University Press.
- The Celtic languages* (1993) ed. Martin J. Ball with James Fife, London: Routledge.
- Comrie, Bernard (1981) *The languages of the Soviet Union*, Cambridge: Cambridge University Press.
- The dialects of Italy* (1997) eds Martin Maiden and Mair Parry, London: Routledge.
- Encyclopedia of the languages of Europe* (1998) ed. Glanville Price, Oxford and Malden: Blackwell.
- Endangered languages* (1991) eds Robert H. Robins and Eugenius M. Uhlenbeck, Oxford and New York: Berg.
- Ethnologue: languages of the world*, 13th edn (1996) ed. Barbara F. Grimes, Dallas TX: Summer Institute of Linguistics.
- FEL II = *Endangered languages: what role for the specialist? Proceedings of the Second FEL Conference* (1998) ed. Nicholas Ostler, Edinburgh: Foundation for Endangered Languages.
- FEL III = *Endangered languages and education: Proceedings of the Third FEL Conference* (1999) ed. Nicholas Ostler, Maynooth: Foundation for Endangered Languages.
- FEL IV = *Endangered languages and literacy: Proceedings of the Fourth FEL Conference* (2000) eds Nicholas Ostler and Blair Rudes, Charlotte: Foundation for Endangered Languages.
- Forbescue, Michael (2005) *Comparative Chukotko-kamchatkan dictionary*, Trends in Linguistics: Documentation 23, Berlin and New York: Mouton de Gruyter.
- Gamkrelidze, Thomas V. and Vjačeslav V. Ivanov (1995) *Indo-European and the Indo-Europeans: a reconstruction and historical analysis of a proto-language and a proto-culture*, Trends in Linguistics: Studies and Monographs 80, Berlin and New York: Mouton de Gruyter.
- The Germanic languages* (1994) eds Ekkehard König and Johan van der Auwera, London: Routledge.
- Germanische Rest- und Trümmersprachen* (1989) ed. Heinrich Beck, *Ergänzungsbände zum Reallexikon der germanischen Altertumskunde* 3, Berlin: de Gruyter.
- Grenoble, Lenore A. and Lindsay J. Whaley (2003) 'The case for dialect continua in Tungusic: plural morphology', in *Current trends in Caucasian, East European and Inner Asian linguistics: papers in honor of Howard I. Aronson*, eds Dee Ann Holisky and Kevin Tuite, Current Issues in Linguistic Theory 246, Amsterdam and Philadelphia: John Benjamins, 97–122.
- Haiman, John and Paola Benincà (1992) *The Rhaeto-Romance languages*, London and New York: Routledge.
- Häkkinen, Kaisa (1984) 'Wäre es schon an der Zeit, den Stammbaum zu fällen?' *Ural-Altäische Jahrbücher*, Neue Folge 4: 1–24.
- Hindley, Reg (1990) *The death of the Irish language*, London and New York: Routledge.
- Janhunen, Juha (1991) 'Ethnic death and survival in the Soviet North', *Journal de la Société Finno-Ougrienne* 83: 111–22.
- (1996) *Manchuria: an ethnic history*, Mémoires de la Société Finno-Ougrienne 222, Helsinki.
- (2004) 'Tungusic: an endangered language family in Northeast Asia', *International Journal of the Sociology of Language* 170: 25–38.
- Jazyki mira: tjurkskie jazyki* (1997) Moskva: Indrik.
- Jazyki mira: mongol'skie jazyki, tunguso-man'čžurskie jazyki, japonskij jazyk, korejskij jazyk* (1997) Moskva: Indrik.

- Jazyki mira: paleoaziatskie jazyki* (1997) Moskva: Indrik.
- Jazyki mira: ural'skie jazyki* (1993) Moskva: Nauka.
- Jazyki narodov SSSR I: Indoeuropejskie jazyki* (1966) general ed. V. V. Vinogradov, Moskva: Nauka.
- Jazyki narodov SSSR II: Tjurkskie jazyki* (1966) general ed. N. A. Baskakov, Moskva: Nauka.
- Jazyki narodov SSSR III: Finno-ugorskie i samodijskie jazyki* (1966) general eds V. I. Lytkin and K. E. Majtinskaja, Moskva: Nauka.
- Jazyki narodov SSSR IV: Iberijsko-kavkazskie jazyki* (1967) general eds E. A. Bokarev and K. V. Lomtatidze, Moskva: Nauka.
- Jazyki narodov SSSR V: Mongol'skie, tunguso-man'čžurskie i paleoaziatskie jazyki* (1968) general ed. P. Ja. Skorik, Leningrad: Nauka.
- Jazyki Rossijskoj Federacii i soseďnix gosudarstv: ěnciklopedija I: A–I* (1997) Moskva: Nauka.
- Jazyki Rossijskoj Federacii i soseďnix gosudarstv: ěnciklopedija II: K–R* (2001) Moskva: Nauka.
- Jazykovaja situacija v Rossijskoj Federacii: 1992* (1992) general eds V. M. Solncev and V. Ju. Mixal'čenko, Moscow [with a detailed English summary: The language situation in the Russian Federation: 1992].
- König, Werner (2001) *dtv-Atlas Deutsche Sprache*, 13. durchgesehene Auflage, München: Deutscher Taschenbuch Verlag.
- Krasnaja kniga jazykov narodov Rossii: ěnciklopedičeskij slovar'-spravočnik* (1994) ed. V. P. Neroznak, Moskva: Academia.
- Lallukka, Seppo (1990) *The East Finnic minorities in the Soviet Union: an appraisal of the erosive trends*, *Annales Academiae Scientiarum Fennicae B* 252, Helsinki: Suomalainen Tiedekatemia.
- Language death and language maintenance: theoretical, practical and descriptive approaches* (2003) eds Mark Janse and Sijmen Tol, *Current Issues in Linguistic Theory* 240, Amsterdam and Philadelphia: John Benjamins.
- Languages in Britain and Ireland* (2000) ed. Glanville Price, Oxford: Blackwell.
- Lectures on language situation: Russia, Estonia, Finland* (2002) ed. Kazuto Matsumura, ICHEL Linguistic Studies 6, Tokyo.
- Lexikon der Romanistischen Linguistik I–VII* (1988–2002) eds Günter Holtus, Michael Metzeltin and Christian Schmitt, Tübingen: Niemeyer.
- Linguistic minorities in Central and Eastern Europe* (1998) eds Christina Bratt Paulston and Donald Peckham, *Multilingual matters* 109, Clevedon: Multilingual Matters.
- Matras, Yarov (2002) *Romani: a linguistic introduction*, Cambridge: Cambridge University Press.
- Minority languages today: a selection from the papers read at the first International conference on minority languages held at Glasgow University from 8 to 13 September 1980* (1981 [2nd edn 1990]) eds Einar Haugen, J. Derrick McClure and Derick Thomson, Edinburgh: Edinburgh University Press.
- Minor languages of Europe: a series of lectures at the University of Bremen, April–July 2000* (2001) ed. Thomas Stolz, *Bochum-Essener Beiträge zur Sprachwandelforschung* 30, Bochum: Brockmeyer.
- Mixed languages: 15 case studies in language intertwining* (1994) eds Peter Bakker and Maarten Mous, *Studies in language and language use* 13, Amsterdam: IFOTT.
- The Mongolic languages* (2003) ed. Juha Janhunen, London and New York: Routledge.
- Nichols, Johanna (2003) 'The Nakh-Daghestanian consonant correspondences', in *Current trends in Caucasian, East European and Inner Asian linguistics: papers in honor of Howard I. Aronson*, eds Dee Ann Holisky and Kevin Tuite, *Current Issues in Linguistic Theory* 246, Amsterdam and Philadelphia: John Benjamins, 207–64.
- Northern minority languages: problems of survival* (1997) eds Hiroshi Shoji and Juha Janhunen, *Senri Ethnological Studies* 44, Osaka: National Museum of Ethnology.
- Philologiae Turcicae Fundamenta I* (1959) eds Jean Deny, Kaare Grønbech, Helmuth Scheel and Zeki Velidi Togan, Wiesbaden: Steiner.

- Podolsky, Baruch (1985a) 'Notes on the Urum language', *Mediterranean Language Review* 2: 99–112.
- (1985b) 'Notes on the Urum (Greek-Tatar) language', *Rocznik orientalistyczny* 44: 59–66.
- Pogarell, Reiner (1983) *Minority languages in Europe: a classified bibliography*, Berlin: Mouton.
- Posner, Rebecca (1996) *The Romance languages*, Cambridge: Cambridge University Press.
- The Romance languages* (1988) eds Martin Harris and Nigel Vincent, London: Croom Helm.
- Rybatzki, Volker (1994) 'Xinjiang im 19. und 20. Jahrhundert', *Journal de la Société Finno-Ougrienne* 85: 149–82.
- Salminen, Tapani (2003) 'Languages, language boundaries and language communities in Europe and beyond', in *I Seminari Internacional de Llengües Minoritàries: ponències i comunicacions = Ist International Seminar on Lesser Used Languages: proceedings*, ed. Angel V. Calpe Climent, Serie Filologica 26, Valencia: Real Academia de Cultura Valenciana, 53–62.
- Schönig, Claus (1997–8) 'A new attempt to classify the Turkic languages' (1–3), *Turkic Languages* 1 (1997): 117–33, 262–77; 2 (1998): 130–51.
- Sibille, Jean (2000) *Les langues régionales*, Dominos 220, Paris: Flammarion.
- The Slavonic languages* (1993) eds Bernard Comrie and Greville G. Corbett, London: Routledge.
- The Turkic languages* (1998) eds Lars Johanson and Éva Á. Csató, London: Routledge.
- The Uralic languages* (1998) ed. Daniel Abondolo, London and New York: Routledge.
- The Uralic languages: description, history and foreign influences* (1988) ed. Denis Sinor, Handbuch der Orientalistik 8: Handbook of Uralic studies 1, Leiden: Brill.
- Vajda, Edward J. (2003) 'Tone and phoneme in Ket', in *Current trends in Caucasian, East European and Inner Asian linguistics: papers in honor of Howard I. Aronson*, eds Dee Ann Holisky and Kevin Tuite, Current Issues in Linguistic Theory 246, Amsterdam and Philadelphia: John Benjamins, 393–418.
- Vakhtin = Vaxtin, Nikolaj (1993) *Korennoe naselenie Krajnego Severa Rossijskoj Federacii*, Sankt-Peterhburg: Izdatel'stvo Evropejskogo Doma.
- Vakhtin = Vaxtin, N. B. (2001) *Jazyki narodov Severa v XX veke: očerki jazykovogo sdviga*, Sankt-Peterhburg: Dmitrij Bulanin.
- Vingt-cinq communautés linguistiques de la France 1: Langues régionales et langues non territorialisées* (1988) Geneviève Vermes, Paris: L'Harmattan.
- The world's major languages* (1987) ed. Bernard Comrie, London and Sydney: Croom Helm.
- Wurm, Stephen A. (1997) 'Two Turkic-based hybrid languages in northwestern China', *Turkic Languages* 1: 241–53.
- Zhào Xiāngrú and Reinhard F. Hahn (1989) 'The Ili Turk people and their language', *Central Asiatic Journal* 33: 260–89.

