

CHAPTER 3

Defining the term "language" and identifying pluricentric languages

The objectives of this chapter

This chapter has two tasks: The first task of this chapter is to define the term “language”. This is done using the example of European languages. The second task is to identify pluricentric languages (PCLs). To achieve this, a solid definition of the term "*language*" is necessary and a prerequisite for the identification and description of PCLs in Europe and elsewhere in the world.

The first step is to define borders and political structures of Europe, as there are different views on where Europe's borders lie. Next, the languages currently used in Europe are described. They are then ranked into different categories in order to identify the PCLs on the continent. In this respect, Europe serves as a model that can be applied to other regions of the world. The pluricentric languages outside Europe are dealt with in a further section.

Part I: Defining the boundaries and countries of Europe

What is the meaning of the term “Europe”?

The continent called “Europe” is a geographical area of around 10 million square kilometres. It borders by the Atlantic Ocean to the west, the Arctic Ocean in the north by and (partly) by the Mediterranean Sea to the south. The Ural Mountains are conventionally regarded as the eastern boundary. Europe includes parts of the Russian Federation even though Uralic and Altaic languages are used on both sides of the mountain range¹. Moreover, this makes Western Kazakhstan (partly) a European country, which until contested by some authors and even the European Union², while Cornell/Engvall (2017) clearly consider it to be a European country³ for different reasons that seem plausible. However, the EU signed a comprehensive treaty in 2015, which is the first of its kind with an Asian country. The country is therefore included in the list of European states. Some authors also delimit Europe at the Black Sea, the Kuma-Manych Depression in Russia⁴, and at the Caspian Sea⁵. This approach would exclude *Turkey, Armenia, Georgia and Azerbaijan*, which are all members of the *Council of Europe*. The latter is not a geographical but a political argument and closes the discussion about the eastern border of the European continent for the time being. In addition, it excludes any attempt to define Europe via so-

called cultural-anthropological arguments, which would restrict Europe to the countries influenced by the Roman-Greek and Christian heritage⁶.

There are also some issues concerning the countries at the western border of Europe. Although *Greenland* is geographically part of the American continent, all countries east of North America are usually considered European. As an autonomous region, Greenland belongs constitutionally to Denmark and can therefore also be included in the list of European territories. The same applies to *Iceland* due to its language and longstanding contacts with (Northern) Europe.

1. The result is the following list of 51 sovereign European countries⁷:

Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Iceland, Ireland, Italy, Kazakhstan, Kosovo⁸, Latvia, Liechtenstein, Lithuania, Luxembourg, North Macedonia, Malta, Moldova, Monaco, Montenegro, Netherlands, Norway, Poland, Portugal, Romania, Russia, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom, Vatican City/Holy See.

2. In addition there are 8 *dependant territories* of European states with a special political (self-governing) status:

- *Akrotiri and Dhekelia*: British overseas territories on the island of Cyprus;
- *Gibraltar*: British overseas territory on the Iberian Peninsula;
- *Island of Guernsey, Island of Jersey, Island of Man*: British Crown territories;
- *Faroe Islands and Greenland*: Constituent countries of the Kingdom of Denmark;
- *Åland*: Self-governing area of Finland;
- *Svalbard / Spitzbergen*: Unincorporated area of Norway.
- *Melilla and Ceuta*: Spanish autonomous cities located on the north coast of Africa and Spanish speaking territories for centuries.

Note (1): The French overseas regions and departments *Guadeloupe, French Guiana, Martinique, Réunion, Mayotte* and the French overseas collectivities *French Polynesia, New Caledonia, Saint Barthélemy, Saint Martin, Saint Pierre and Miquelon, Wallis and Futuna*, which have their own statutory laws are not included here in the list of dependant territories as they are all situated outside the European continent. To include them would mean to include many very different languages in the list of European languages, which would overstretch the term “European language”.

Note (2): The same approach is applied to the constituent parts of the Kingdom of Netherlands, which are all situated in the Caribbean Sea: *Bonaire,*

Saint Eustatius, Saba, Aruba, Curaçao, Saint Maarten. There is just one typologically non-European-language – *Papiamentu* – (a Creole language), which is spoken in these territories (but not included either in the list of European languages.)

3. Finally, there are 10 *disputed territories* which are in some cases de facto independent states or in other cases incorporated areas in other polities with little or no international recognition¹⁰:

- *Abkhazia and South Ossetia*: Breakaway territories of Georgia which declared independence;
- *Autonomous Republic of Crimea*: Occupied area of Ukraine by Russia;
- *Republic of Artsakh / Nagorno-Karabakh Republic*: Breakaway territory of Azerbaijan, formally independent republic only supported by Armenia and dissolved after the war in 2023;
- *Luhansk People's Republic and Donetsk People's Republic*: Breakaway territories of Ukraine, which are presently at war with Ukraine and now occupied by Russia since 2022;
- *Northern Cyprus*: Breakaway territory of the Republic of Cyprus, occupied by Turkey;
- *Transnistria*: Breakaway territory of the Republic of Moldova.

In all, there are 82 European political entities that form the basis of the next step: listing the languages used on European territory and then sorting out those which can be considered pluricentric.

Part II: Defining the term “language”, setting up a complete list of European languages and a list of European pluricentric languages

1. Problems in defining and listing the European languages

1.1 Existing lists denoting “European languages”

The available literature dealing with the topic “European language” (EL) shows a confusing variety of names and provides numbers of languages used in Europe that vary significantly. Furthermore, there is a significant lack of literature on the subject, illustrated by the fact that Encyclopaedia Britannica does not have such an entry (Wikipedia has a section called “Languages of Europe” with differences between language editions). Six main sources were identified in the search for relevant literature:

(a) The “*Wieser Enzyklopädie: Lexikon der Sprachen des Europäischen Ostens und Westens. Drei Bände*” [Wieser Encyclopedia: Lexicon of the Languages of the European East and West. Three Volumes]: This comprehensive work in three

volumes, edited by different authors presents on 2.300 pages a total of 145 living European languages (and also some extinct languages), many of which are unique to this encyclopaedia. It is surely the most extensive publication in the field, in which expert linguists are involved as authors and editors: It comprises Vol. 1-2 (2008) “*Sprachen des Europäischen Westens*” [Western European Languages] and Vol. 3 (2002) “*Lexikon der Sprachen des europäischen Ostens*” [Lexicon of the Languages of the European East].

(b) “*Das Sprachenmosaik Europas*” [The language mosaic of Europe] by Harald Haarman (2011) published on EGO “*Europäische Geschichte Online*”¹¹ [European History Online]. This list comprises 97 languages and is comparatively short.

(c) The list of “*European Languages*” published in three editions of *Wikipedia*: The English, French and Spanish versions.

(d) The list of languages protected by the *European Charter of Minority Languages*, which comprises 91 languages¹². This list is included because of its reliability and in order to obtain a complete list of all European languages. Another reason is that many so called “minority languages” are sometimes more accurately described as regional languages (e.g. Hungarian) and therefore qualify as national varieties.

(e) The list of official languages of the countries of the European Union and the other European states listed above. It must be noted at this point that the term “official language” is ambiguous and a source of uncertainty as this concept differs from country to country.

2.1 Problems with the definition of the terms “language” and “European language”

(a) Different numbers of languages in different lists

None of the three major sources above shows the same number of languages. This is best illustrated by a comparison between three major language specific editions of *Wikipedia*: *Wikipedia* English lists 121 languages, *Wikipedia* French 108, *Wikipedia* Spanish 132. The differences arise – to show just an example – as the Spanish list includes 32 Caucasian languages and 10 varieties of Romani that are not found in other language lists. This results in large differences in the number and names of languages in the lists. Some languages are found in all lists and some in only four, three, two or one, which required extensive research in additional sources.

(b) Different names with different spellings for the “same” language

A second, rather confusing problem was that a number of languages have different names, different spellings and as such can be found in the lists consulted. Examples for this are:

Abkhaz: Abxaz, Abkhazian;	Mari: Cheremis; Kabardian: Kabardino-Cherkess, East Circassian;
Aramaic: Aisor, Neo-Assyrian, Neo-Aramaic	Lezgian: Lezgi/Lezgin;
Arbëresh: Arbëneshtë, Arbanasi Albanian	Mirandese: Asturleonese, Mirandés
French Channel Islands: Jèrriais. Jèrriais / Dgèrnésiais)	Rhaeto-Romance: Rumantsch Grischun, Romansh, Rumantsch etc.

(c) Unclear concepts related to the term “language”

A third and by far the most significant complication arose from the fact that there are language names in the lists (e.g. the English Wikipedia) like “*Austro-Bavarian*”, “*Swabian German*” [Schwäbisch], “*Mainfränkisch*“ [Main Franconian], “*Ripuarian (Platt)*” and “*Palatinate German*” [Pfälzisch] which are traditionally regarded as regional varieties of German and are usually listed as (historical) “dialects” of German. Linguistically and historically they may be considered as “languages” but do not function today as such; they are used as spoken everyday language and generally not in writing. The list also contains language names like “*Walloon*” [Belgium French], “*Silesian*” [Schlesisch] and “*Silesian German*” [Schlesisches Deutsch], which are different language types: national varieties (Belgium-French), nearly extinct dialects (Silesian German) and mixed languages (Silesian). They are all listed under the heading of “European language”.

This is quite revealing, as it involuntarily provides insight into one of the central problems of contemporary linguistics and language policy: there is no consensus about the most fundamental term of linguistics: “*language*” – and thus also not on derived terms such as “*pluricentric language*”, “*official language*” and related terms like “*national variety*”, “*dialect*”, “*minority language*” and other terms denoting status types of languages. This reveals a central source of conflict in PCLs and beyond, and a massive shortcoming of the linguistic sciences.

(d) A new approach to define the pluricentric languages of Europe and their national varieties: Defining the term “language” by classifying “lingoids” according to their status in order to become “languages”

The confusing situation around the term “*language*” and “*European lan-*

guage” led me to compile a complete list of all languages included in the five lists (plus consulting further online resources), to classify them according to their status in order to be able to define central terms of variational linguistics and terms underlying the concept of PCLs.

The aim of this section is to identify the pluricentric languages of Europe. A first attempt made in this direction was my paper published in 2003¹³. This paper, however, is outdated today due to developments in the theory of pluricentricity and changes in the political situation in Europe. A new approach is needed that includes the definition of basic linguistic concepts like “language”.

(e) Criteria and concepts to define the term “language”

In this chapter, a list of “European languages” is presented. To achieve this the term “*language*” had to be defined at first. The term “*lingoid*” is introduced for this purpose.

The term “*lingoid*” refers to any means of human communication that is traditionally labelled as “*language*”, but whose status in the social sphere is not yet decided. It is a neutral hyperonym for any kind of unclassified “*language*”. A “*lingoid*” becomes a “*language*” if this communication system has achieved political/social acknowledgement and an official function in a given social/political realm it is attributed to.

The reason for the introduction of this term is that many ambiguous and unclear *lingonyms* (language names) emerged, when the list of European languages was compiled.

It was not clear what type of “*language*” is meant by a series of language names or whether they “*existed*” at all, especially when they appeared only once in the lists. It seemed therefore reasonable to have a *neutral umbrella term* in analogy to the term “*phone*” in phonology that describes any vocal utterance on the level of phonetics whose phonological status is not (yet) decided.

- The following classification of European *lingoids* is based on the *criterion of recognition by a political or societal authority that has the power to decide over the status of lingoid(s) being used on the territory the authority governs*.
- The “*political or societal authority*” would be in the most cases the political institutions of a country (government of a nation or of a state/department/region, parliament etc.). Outside the formal structures of nations, the body passing such decisions could be a council of elders, one or more leaders of a social or ethnic group etc.
- These political entities and powers decide whether a given “*lingoid*” “*exists*” (on the territory), which means that it is given a commonly used name

and whether and how it may / must be used in this realm.

- It is the political and social process that transforms a “*lingoid*” into a “*language*”, and it is the political entity that endows “*lingoids*” with rights and brings them into existence as “*languages*”.
- A “*language*” in this concept is *therefore a lingoid that has full recognition and an ascribed status by a political entity*.
- *Lingoids without this recognition exist of course but they cannot be called a “language” in the sense described here¹⁴. Based on this it is possible to define the term “European language” (and later on the term “European pluricentric language”).*

Definition: A European language is a lingoid that is native to one of the (presently) 82 political entities of the European continent, used by speakers living there, recognized by a political entity and endowed a specific status that defines its use and its rights within the given realm.

Part III: A comprehensive list of all European languages and lingoids - Classified according to their occurrence and status

The number of European lingoids that are potentially “languages”

The survey described above identified 238 *European lingoids*, of which 199 are *languages according to the previous definition and 39 are unrecognized minority languages or languages of migrant communities (mostly with very few speakers) or regional languages (often referred to as “dialects”) or languages used by itinerant communities that are without a formal status and do not formally exist*. The number of lingoids/languages found might in fact be larger, as there are smaller lingoids in the Balkans and in Russian Federation that are probably not included in any of the lists consulted.

Category (1): Recognised lingoids with predominant rights that are used with maximum status in only one nation: *National monocentric languages* (21):

Features:

- They encompass a well-defined territory where they are a source of social cohesion;
- They have the highest social status in this area. The status is conferred on the lingoid by state institutions and usually secured by laws;
- Their use is prescribed for certain text types and communicative situations (administration, institutions of the state etc.);

- Are part of identity and self-identification of the nation and its inhabitants;
- These lingoids are themselves norm-setting centres, and also function as a model for the same language in other nations with the status of a minority language.

List: *Belarusian, Bosnian, Bulgarian, Czech, Danish, Estonian, Finnish, Georgian, Icelandic, Latvian, Letzebuergesch (Luxembourgian), Lithuanian, Macedonian, Maltese, Montenegrin, Polish, Rumantsch Grischun, Slovakian, Slovenian, Ukrainian, Welsh (y Gymraeg).*

Note (1): Welsh (y Gymraeg) is only recognized as a “minority language” but is the second official language of Wales, which is constitutionally part of the United Kingdom and legally a sovereign state within the United Kingdom. The language is also an important source of identity for the population of Wales and unique to this political entity. It can therefore be considered a co-official national monocentric language.

Note (2): Rumantsch Grischun is one of four national languages of Switzerland and limited to this country. This language must therefore also be categorised as a monocentric national language.

Category (2): Recognised lingoids that are used within more than one nation in which they have the highest status: *National pluricentric languages* (27):

Features:

- Encompass a well-defined territory in at least two nations in which they are a source of social solidarity. Several nations/self-governing regions share this language, to a greater or lesser extent;
- They have the highest status - conferred by political institutions and generally guaranteed by law;
- Their use is prescribed for certain text types and communicative situations (administration, institutions of the state etc.);
- The lingoid is part of the identity and self-identification of the nation and its inhabitants through speakers who are aware of their own linguistic and pragmatic characteristics;
- The participating nations have a sufficiently large number of speakers to establish their own linguistic centre, which codifies and standardises the variety after a longer period of time.

List: *Albanian, Armenian*, Azerbaijani North (Azeri), Basque, Catalan, Croatian, Danish, Dutch, English, French, Frisian*, German, Greek, Hungarian*, Irish (Ireland*

Gaelic), Italian, Norwegian, Occitan, Portuguese, Romanian, Russian, Scots, Sami*, Serbian, Spanish (Castilian), Swedish, Turkish*.*

Note (1): There are several languages in this list with a high degree of internal fragmentation, which makes their status as PCL questionable: *Frisian, Occitan, and Sami*. They are listed here until their status as PCL is clarified. They will be dealt with in detail in Chapter (9)

Note (2): Specific features of this type: The term “*pluricentric language*” does NOT refer to a single lingoid. It is an umbrella term for several (at least two) acknowledged national lingoids (national varieties) which in their territory have (in most cases) the highest status and act like monocentric national languages.

Note (3): It is important to note that a substantial number of speakers at least of the quality of a regional language is necessary to develop a norm-setting centre and thus a distinct variety of its own, which is codified and standardized after some time.

Category (3): Recognised lingoids (and thus languages) used in just one nation – but with a common history and shared linguistic features with the language(s) of other nations: *National varieties of pluricentric languages*: (65):

Features:

- Encompass a well-defined territory in a nation or region in which they are a source of social solidarity. Other nations/self-governing regions share this language;
- They have the highest status - conferred by political institutions and generally guaranteed by law;
- The lingoid is part of the identity and self-identification of the nation and its inhabitants through speakers who are aware of their own linguistic and pragmatic characteristics;
- Are languages shared with (an)other nation(s)
- These lingoids usually do not have an official language name of their own – they are just named after the name of language type (2).
- They often share language regulation bodies that regulate orthography and other aspects of the formal written language (written standard language)
- They are called “*national varieties*” and are the realization of a pluricentric language in a specific nation/region.

- The participating nations have a sufficiently large number of speakers to establish their own linguistic centre, which codifies and standardises the variety after a longer period of time.

Note (1): The marking of a language name with an asterisk (*) in this list is an indication that there is some particularity with this NV, which means that there is a deviation in official status or issues about the language name etc. These issues will be discussed further down in this chapter.

Note (2): The national varieties *constitute* the pluricentric language. They must therefore be the primary source of description as a whole as they represent single nations and have the same status in each nation like a monocentric language.

Note (3): In the following list, the name of the NVs of PCLs unusually composed of the name of the language in the first position, followed by the name of the country in second position. This system was chosen here exclusively in order to group all national varieties of single PCLs together in one place. The standard form of naming NVs of PCLs it is exactly the other way round.

List: *Albanian Albanian, Albanian Kosovo, Albanian North Macedonia, Armenian Armenia, Armenian Artsakh*, Armenian Western**; *Azerbaijani (Azeri) Azerbaijan, Azerbaijani (Azeri) Russia, Basque Spain, Basque France, Catalan Andorra, Catalan Balears; Catalan Catalunya, Catalan Valencia, Croatian Bosnia Herzegovina, Croatian Croatia; Dutch Belgium (Flemish), Dutch Netherlands, English England*, English Ireland, English Northern Ireland*, English Malta, English Scotland*, English Wales, French Belgium, French France, French Monaco, French Switzerland, German Austria, German Belgium, German Germany, German Italy; German Liechtenstein, German Luxembourg, German Switzerland, Greek Cyprus, Greek Greece, Hungarian Romania*, Hungarian Serbia*, Hungarian Slovakia*, Hungarian Ukraine*, Irish Ireland, Irish Northern Ireland, Italian Italy, Italian San Marino, Italian Switzerland, Norwegian (Bokmål), Norwegian (Nynorsk), Occitan Spain, Occitan France, Portuguese Portugal, Romanian Romania, Romanian Moldova*, Russian Estonia*, Russian Latvia, Russian Lithuania, Russian Ukraine*, Sami Norway*, Sami Sweden*, Sami Finland*, Scots Scotland, Scots Ulster, Serbian Serbia, Serbian Bosnia Herzegovina, Spanish Spain, Swedish Finland, Swedish Sweden, Turkish Cyprus, Turkish Turkey.*

Category (4): Recognised lingoids used in sections of the territory of a certain nation: *Official regional languages (66 *)*

Features:

- They occupy a well defined but limited section of a nation's territory alongside the national lingoid(s);
- They have a status that is restricted to this region with legally guaranteed rights;
- The right to use these regional languages in certain official text types and communicative situations (administration, institutions etc.) is limited to this area;
- They are part of the identity of the region and its inhabitants and often take their name from the region they encompass.
- If there is a considerable number of speakers, the regional language can become a norm-setting centre.

List: Abaza (*Adyghe*), Abkhaz (*Abxaz, Abkhazian*), Adyghe (*West Circassian*), Aghul, Aragonese (*Spain*), Aramaic / Aisor (*Neo-Assyrian, Neo-Aramaic*) (*Kurdish Autonomy Region*), Aranese (*Occtitan*), Archi, Asturian (*Bable*), Avar (*Avaric*), Bashkir, Basque, Catalan, Chechen, Chuvash, Corsican, Dargwa, Erzya, Farese (*Faroese*), Franco-Provençal (*Arpitan*), Frisian-North, Frisian-West, Friulian, Friulian (*Furlan*), Gagauz (*Gagausian*), Galician (*Galego*), Gallurese, German-South-Tyrol, Greenlandic (*Inuit*), Ingush, Istro-Venetian (*VENET*), Judeo-Tat (*Juhuri*), Kabardian (*Kabardino-Cherkess, East Circassian*), Kalmyk Oirat, Karachay-Balkar (*Karatschai-Balkarian*), Komi-Permyak, Komi-Zyrian, Kumyk, Ladinian (*Dolomite Ladin*), Lak, Leonese (*Leonesian*), Lezgian (*Lezgi/Lezgin*), Limburgish, Low German (*Plattdütsch, Lower Saxon*), Mari (*Cheremis*), Mirandese (*Asturleonese, Mirandés*), Mòcheno (*Fersentalerisch*), Mordvinian Erzya, Mordvinian Moksha, Nenets (*Yurak*), Nganasan, Nogai (*Nogay, Nogai Tatar*), Occitan, Ossetian (*Ossete, Ossetic*), Piedmontese, Rutul, Northern Sami, Sardinian, Sassarese, Sicilian (*Siculo, Calabro-Sicilian*), Tabasaran, Tat / Tati, Tatar Crimean (*Crimean*), Tatar Kazan, Tsakhur (*Tsaxur, Caxur*), Udmurt (*Wotjakian*), Venetian.

Note (1): A large number of the regional languages of this list are official languages of the federal republics of the Russian Federation, which according to article 68 of the Russian constitution have the right to their own state language other than Russian.¹⁵ However, a new education bill passed in 2018 makes the teaching of the languages of the republics optional and reduces them to two hours per week. This has seriously undermined the right to use

the federal regional languages and aims to Russify the Russian Federation.¹⁶

Note (2): Many lingoids of this type would otherwise often be called “*dialects*” or “*regional varieties*” of a national language as it is the case with German, for example. However, in many other cases (as in Russia) they are “*regional languages*”, because they are recognised on the level of republics or regions and have an official function in this area. This is particularly the case in Spain and Italy, where there is regional devolution, which includes the official recognition of the languages of these regions.

Category 5: Recognised lingoids different from national language(s) and only used in a small part of a certain nation and by a small minority of its inhabitants with very reduced rights: *Protected minority languages*¹⁷ (57 / 93)¹⁸:

Features:

- In addition to the national lingoid(s), they usually comprise well-defined defined parts of a nation's territory;
- They have a restricted status and only have linguistic rights for a specific social group, which are granted to them by the state and usually guaranteed by law. They "exist" because they have been granted some rights in a limited area.
- Their right of use in certain official text types and communicative situations (administration, institutions, etc.) is restricted to this area/group;
- These lingoids are often also used in other nations, especially in neighbouring countries with an adjacent contiguous language area.

List: *Arabic Cypriot Maronite (Kormakiti Arabic), Arbëresh (Arbëneshtë, Arbanasi Albanian), Aromanian (Macedo-Romanian / Vlach), Beás (Boyash Roma, Beash, Rudari, Ludari), Bunjevac (Bunjevački govor, Bunjevački jezik), Cimbrian (Northern Italy), Cornish (Kernowek), Crimean Tatar, Croatian Moravian, Finnish Kven, Finnish-Meänkieli (Tornedalsfinska), French Channel Islands (Jèrriais. Jèrriais / Dgèrnésiais), Frisian-Sater, Gaelic Manx, Gaelic-Scottish (Scots Gaelic), Gallo, Griko (Grico, Italiot Greek), Hebrew, Istriot (Istro-Romanian, Vlach), Italo Albanian (Arbëresh), Judaeo-Spanish (Ladino), Karaim (Karaimic), Kashubian (Kashub), Krymchak (Krimchak, Judeo-Crimean Tatar), Lemko (Lemkian, Ruthenian, Rusyn Poland), Molise Slavic (Slavic in Southern Italy), Romanes (Romani, Romani Chib, Roma language), Romani Sinti (Sintenghero Tschib(en), Sintitikes, (Romanes), Romani Vlach, Rusyn (Ruthenian, Ruthene), Sami Inarisami, Sami Lulesami, Sami-Skolt, Sami / Eastern Skolt Sami, Sami-South Sami, Sorbian Lower,*

Sorbian Upper, Tatar Finland (Mishar), Vepsian (Veps), Walachian, Yenish (Yeniche, Jenisch), Yiddish.

Note (1): The first figure “57” above indicates the number of protected minority languages (PMLs) that are unique in Europe, while the second number “93” indicates the number of PMLs listed in the European Charta of Minority languages across all member states of the Council of Europe (some minority languages are protected in more than one country).

Note (2): Many lingoids of this type often only have a small number of speakers. Due to the small number of speakers and lack of resources, in most cases they are not norm-setting centres. In addition, these lingoids often differ substantially in their linguistic substance from the mother-variety, which often leads to diglossia.

Category (6): Recognised lingoids derived from national language(s) that are only used by specific groups of inhabitants of nations with specific linguistic needs because of their physical impairments: *Protected sign languages as a type of minority languages (31 / 20)*¹⁹:

Features:

- They share the same national territory like the monocentric national languages (Type 1) or the national varieties of pluricentric languages (Type 3).
- Are the primary means of communication of specific groups of the population with physical impairments (hearing impairment, muteness etc.).
- Have a reduced status and have linguistic rights only for a specific social group – which is bestowed by the state and usually secured by law.

List: Type (1) Monocentric sign languages (31)

Armenian SL, Azerbaijani SL, Bosnian SL, Bulgarian SL, Croatian SL, Czech SL, Danish SL, Estonian SL, Finnish SL, Georgian SL, Hungarian SL, Icelandic SL, Irish SL, Latvian SL, Lithuanian SL, Macedonian SL, Maltese SL, Moldavian SL, Northern Ireland SL, Norwegian SL, Polish SL, Portuguese SL, Romanian SL, Russian SL, Serbian SL, Slovakian SL, Slovenian SL, Spanish SL, Turkish SL, Ukrainian SL, Yugoslav SL

Type (2) Pluricentric sign languages (20)

Albanian-Albania SL, Albanian-Kosovo SL, Catalan Catalonia SL, Catalan-Valencia SL, Dutch-Flemish-Belgium SL, Dutch-Netherlands- SL, English-British SL, English-Northern Ireland SL, French-Belgium-Walloon SL, French-France SL, French-Switzerland SL, German-Austrian SL, German-Germany SL, German-

Switzerland SL, Greek-Cyprus SL, Greek-Greece SL, Italian-Italian SL, Italian-Switzerland SL, Swedish-Finland SL, Swedish-Sweden SL.

Category (7/8): Lingoids of type (4) and (5) with no recognition by the political entities where they are occurring and thus only existing informally and without explicit rights (16/21): *Unrecognised and unprotected regional / minority languages - - Invisible languages (I):*

Features:

- Its speakers usually encompass only small sections of the territory of a nation or are scattered about the whole nation. They may be of the type (4) regional or (5) minority languages;
- Have no recognition and therefore have no status and do not officially “exist”;
- Only in very few cases, they develop actions to standardise and codify their language (if the social group is powerful enough).
- This situation leaves the specific social groups without any right of use in official national contexts;

List (6): Following type (4) – unrecognised/unprotected regional languages:

Asturian, Breton, Chakavian (Kajkavian), Extremaduran, Gascon (Aranés, Aranese, Arnais, Aranese Occitan), Karelian Proper, Karelian Olonets, Ligurian, Lombard, Masurian, Mingrelian (Megrelian, Iverian), Neapolitan, Resian (Slovenian in Italy), Shtokavian, Silesian, Svan, Udi, Vičsch, Vöro, Walachian, West Polesian, Wymysorys (Vilamovian) etc.

List (7): Following type (5) – unrecognised/unprotected minority languages:

Arvanitika Albanian (Arvanitic), Bats (C'ova-Tuš, Batsi), Budukh, Domari (Middle Eastern Romani, Tsigene, Luti, Mehtar), Judeo-Italian (Italkian), Kryts (Kryc), Kurdish Northern (Kurmanj), Megleno-Romanian (Meglenitic, Moglenitic), Pomak (Greek Bulgarian), Sami-Kildin Sami (Kola Sami), Sami-Pite Sami, Scottish Cant (Scottish Traveler Cant), Talysh, Urum, Votic etc.

Note: The list above is probably not complete, as some of the lingoids belonging to this category are not mentioned in the lists that served as the basis for this compilation or are described in literature that is difficult to access.

Category (8): Nationless lingoids (diaspora languages) – migrant lingoids with no recognition by political entities: (Partly) Invisible languages (II):

- Category (8) includes lingoids that are used by traditional diaspora commu-

nities, which came into existence due to migration in the course of political events or due to labour migration.

- These lingoids share all features of categories (6) and (7) and have no rights whatsoever. This is particularly the case with diaspora languages that came into existence through labour migration and are not recognized in any European nation. Some traditional diaspora languages like *Yiddish* or *Armenian* sometimes happen to achieve some recognition in some countries as minority language (see above category (5)).

List (only to name some as an example):

Assyrian (Asidor, Neo-Aramaic), Hebrew, Judeo-Spanish, Kurdish, Yiddish, Turkish, Romanian, Polish, the languages of Ex-Yugoslavia etc. in Austria, Germany, Switzerland and many other countries etc.

Category (9): Nationless lingoids of the itinerant people and territorially unbound languages with no recognition by political entities: Invisible languages (III):

- Category (9) includes lingoids that are mainly used by itinerant people who travel around in different countries.
- These are the most powerless lingoids as they “own” no explicit territory, no rights, and no nation of origin and thus have no support. Some of these lingoids are also secret languages (crypto languages) used solely for communication within the community.
- Some of these lingoids are recognized in some European countries as minority languages (Switzerland, Scotland, UK).

Examples:

Varieties of Romanes, Yenish, Polari (Palari, Palarie), Scottish Cant (Scottish Traveller Cant), Shelta (Cant, Sheldru).

Footnotes:

¹ Haarman (2011)

² See: Euroactiv: EU hesitates to recognise Kazakhstan as European [06.11.2017].
[<https://www.euractiv.com/section/central-asia/news/europe-hesitates-to-recognise-kazakhstan-as-european/>] [acc. 17.11.2018]

³ Cornell/Engvall (2017)

⁴ <https://www.britannica.com/place/Kuma-Manych-Depression> [acc. 06.01.2025]

⁵ See: <https://www.britannica.com/place/Europe> (acc. 06.01.2025)

⁶ See Grzega (2006)

⁷ Salewski (2004) and Wikipedia:

https://en.wikipedia.org/wiki/List_of_sovereign_states_and_dependent_territories_in_Europe (acc.

-
- 06.01.2025)
- 8** As of 03/2025 Kosovo is only recognised by 104 countries of the United Nations. It is considered an independent state here, as it has all features of a functioning statehood.
- 9** Just to give one example: The regional languages of French Guiana “include French Guianese Creole (not to be confused with Guyanese Creole), six Amerindian languages (Arawak, Palijur, Kali'na, Wayana, Wayampi, Emerillon), four Maroon Creole languages (Saramaka, Paramaccan, Aluku, Ndyuka), as well as Hmong Njua. Other languages spoken include Portuguese, Hakka, Haitian Creole, Spanish, Dutch, English, and Tamil, and Caribbean Hindustani” (Quoted from Wikipedia English edition: French Guiana).
- 10** As of 01.11.2025: Kosovo: Recognised by 104 out of 192 United Nations member states, Abkhazia and South Ossetia: 5, Republic of Artsah (Nagorno-Karabakh Republic): 1 (now defunct), Northern Cyprus: 1, South Ossetia: 5, Transistria: 0.
- 11** <http://ieg-ego.eu/de/threads/crossroads/sprachenmosaik/harald-haarmann-das-sprachenmosaik-europas> [acc. 15.02.2025]
- 12** <https://www.coe.int/en/web/european-charter-regional-or-minority-languages/languages-covered> (acc. 10.02.2025)
- 13** Muhr (2003): Die plurizentrischen Sprachen Europas. [The pluricentric languages of Europe]
- 14** A good example for the validity of this concept is my home town Graz in Austria. About 150 lingoids (languages) are spoken in a town of not more than 300,000 inhabitants. However, there is only one language in the sense depicted above: Austrian German. It is the only language that is / and may be officially used in the public domain (administration, media, schools etc.) and is also used by most inhabitants in their private lives. All other “languages” are restricted to the private domain and (apart from publications of engaged linguists) do not exist officially.
- 15** https://en.wikisource.org/wiki/Constitution_of_Russia [accessed 25.02.2025]
- 16** <https://www.idelreal.org/a/28937644.html>; <https://theconversation.com/russia-is-cracking-down-on-minority-languages-but-a-resistance-movement-is-growing-101493> [accessed 25.02.2025]
- 17** For this type see the European Charter of Minority languages: <https://www.coe.int/en/web/european-charter-regional-or-minority-languages/languages-covered> [acc. 25.02.2025]
- 18** The first number (57) names minority languages that are specific to a single country, the second number (93) counts all minority languages – many of them exist in several countries.
- 19** The first number (30) indicates the number of monocentric sign languages, the second (20) the number of pluricentric sign languages. The list is based on the article „List of sign languages“ on Wikipedia (English edition).