Elisabeth Piirainen

"Widespread idioms in Europe and beyond": Lesser-used languages and their contribution to a "Lexicon of Common Figurative Units"

1. Preliminary Remarks: Common idioms in European languages

According to one line of early phraseology research, idioms were considered to be a highly distinctive part, if not the innermost part of a language, which led to the idea that idioms of one language had no parallels in the idioms of other languages and that they were ultimately untranslatable. The idea that the figurative lexicon of a given language provides the basis for a worldview that mirrors some national-cultural character and national mentality originated in national romantic thinking, which thought of languages as being identical to nations or cultural communities – something that even in Europe alone is nowhere the case. This idea has largely been disproved and abandoned since. More recent studies have occasionally pointed to the fact that there are indeed extensive similarities among the idioms of several European languages. These assumptions are fully confirmed by the results of the large-scale international project "Widespread idioms in Europe and beyond".[1]

The project started in 2005 and is still ongoing. One of the goals is the systematic investigation into idioms that exist in many European languages (and beyond) in about the same lexical structure and figurative core meaning. First, we had to develop a theoretical framework and a suitable metalanguage. The next objective was to systematically discover *widespread idioms* across as many languages as possible. Data were gathered for 74 European and 19 non-European languages.[2] With the help of more than 300 respondents who patiently completed the many questionnaires for their native languages we were able to identify about 500 idioms which are common in a large number of languages, even in geographically, genetically and culturally distant languages, i.e. idioms that form the core of a "Lexicon of Common Figurative Units". All results came as a real surprise. We now know which Europe-wide common idioms actually exist: by far not all of them are of ancient or biblical origin, as is often assumed.

Among the European languages we analyzed there are major and minor languages (see Section 2). In this article, we want to look at the lesser-used languages included in the project: Are there differences between the major languages, i.e. the well-established literary languages, often also national languages, and the minor languages in the realm of the common figurative lexicon? Are the smaller languages involved in constituting the uniformity of idioms of European languages, and to what extent? What could be the reasons for the differences? Before we look at some examples (Section 4), the languages of our project and the term *lesser-used language* should be considered in more detail

---

[1] First results have been published in (PIIRAINEN 2012); a second Volume of a "Lexicon of Common Figurative Units" is in the works.

[2] We had no access to a further 70 languages spoken in Europe, i.e. the many declining or seriously endangered lesser-used languages of Eastern and Southeastern Europe, among them mainly Finno-Ugric and Caucasian languages.

(Section 2), and the problems we have in collecting idioms of lesser-used languages should be discussed (Section 3).

2. Major languages and lesser-used varieties represented in the project: an overview

The sociolinguistic status of the languages of Europe varies greatly between official, co-official, lesser-used, declining, etc. A division into two main groups becomes apparent: It is the sociolinguistic division between the larger well-established literary languages and the minor and minority languages.[3] This division is clearly reflected in the project "Widespread idioms". Off the 74 European languages we analyzed approximately one half belong to the first group. Language varieties of the second group can be roughly subsumed under the term *lesser-used languages* because they share some of the following features:

(i) The language is quite isolated, restricted to a small area with only a few speakers.
(ii) The language is less vital or declining: the number of speakers is decreasing.
(iii) The linguistic competence of the speakers has declined; only representatives of the oldest generation still speak the language fluently. The inter-generational transmission is restricted.
(iv) The number of domains is limited, e.g. language use is restricted to a few mostly private and informal domains.
(v) There is a direct competition to a large, more prestigious language. This is often the official and overarching national language.
(vi) The language lacks a written register or traditions of writing and literature.
(vii) The language lacks a higher degree of standardization; it is used predominantly in oral form.

There is not a minimum number of speakers among the criteria. In terms of numbers, Icelandic with its only 310,000 speakers could be considered a "small language"; however, none of the features above applies to Icelandic. The language is used in all domains almost without competition. Similar assessments can be made for the other languages spoken on the islands, for Maltese and the much smaller Faroese. In contrast, several languages that currently have about half a million (e.g. Aromanian, Breton, Mari, Udmurt) or even 1.5 million speakers (Sardinian) can be viewed as lesser-used and endangered because the transmission to new generations is not guaranteed. A few languages of our project cannot be clearly assigned to one of both groups: Luxembourgish, Yiddish, Catalan, Galician and Tatar are "borderline cases".

Here follows a brief overview of the languages included in the project and their sociolinguistic status, starting with the lesser-used languages. For the sake of clarity, we group the languages according to the language families they belong to: (i) Indo-European, (ii) Finno-Ugric, (iii) Turkic, (iv) Semitic, (v) Caucasian and (vi) Basque.

---

[3] For various notions of *minor language* and *minority languages* see SHERZER/STOLZ (2003), among others. For definitions of the term *endangered language* with its gradations such as *seriously endangered*, *irreversibly endangered*, *near-extinct* cf. Crystal (2000), NETTLE/ROMAINE (2000), HARRISON (2007, 2010), IDSTRÖM/PIIRAINEN (2012) and the UNESCO Atlas of the World's Languages in Danger.

(i) Indo-European languages in Europe:
(a) The 15 **Germanic** languages of the project consist of five lesser used languages (*Scots*, *North Frisian*, *West Frisian*, *Low German*, *Swiss German*), eight major languages (*Icelandic*, *Faroese*, *Norwegian*, *Swedish*, *Danish*, *English*, *Dutch*, *German*, *Luxembourgish*), as well as *Luxembourgish* and *Yiddish* whose status cannot be clearly evaluated.
(b) All five **Celtic** languages are lesser-used and partly declining varieties: *Irish*, *Scottish Gaelic*, *Cornish*, *Welsh* and *Breton.*
(c) Among the 16 **Romance** languages there are nine lesser-used or minority languages (*Provençal*, *Occitan*, *Ladin*, *Romansh*, *Francoprovençal*, *Venetian*, *Sardinian*, *Mirandese*, *Aromanian*) and five major languages (*French*, *Italian*, *Spanish*, *Portuguese*, *Romanian*). The status of *Catalan* and *Galician* is not clear; currently, they are developing languages.
(d) Both of the **Baltic** languages, *Latvian* and *Lithuanian*, are major languages.
(e) The 14 **Slavonic** languages consist of two lesser-used languages (*Kashubian* and *Sorbian*, including the varieties *Upper* and *Lower Sorbian*) and twelve major languages (*Russian*, *Belorussian*, *Ukrainian*, *Czech*, *Slovak*, *Polish*, *Slovene*, *Croatian*, *Bosnian*, *Serbian*, *Macedonian*, *and Bulgarian*).
(f) **Albanian**, an isolated Indo-European major language.
(g) **Greek**, an isolated Indo-European major language.
(h) **Armenian**, an isolated Indo-European major language.

(ii) Finno-Ugric languages in Europe:
(a) *Hungarian*, the only **Ugric** language spoken in Europe, is a major language.
(b) Among the four **North-Finnic** languages there are two declining languages (*Karelian* and *Veps*) and two major languages (*Finnish* and *Estonian*).
(c) Both *Udmurt* and *Komi*, the two **Permic** languages represented in the project, are under pressure from Russian and belong to the lesser-used endangered languages.
(d) The three **Volgaic** languages *Mari*, *Moksha Mordvin* and *Erzya Mordvin* are also endangered languages of oppressed minorities.
(e) The **Saamic** languages, represented by *Inari Saami* in our project, are seriously endangered.

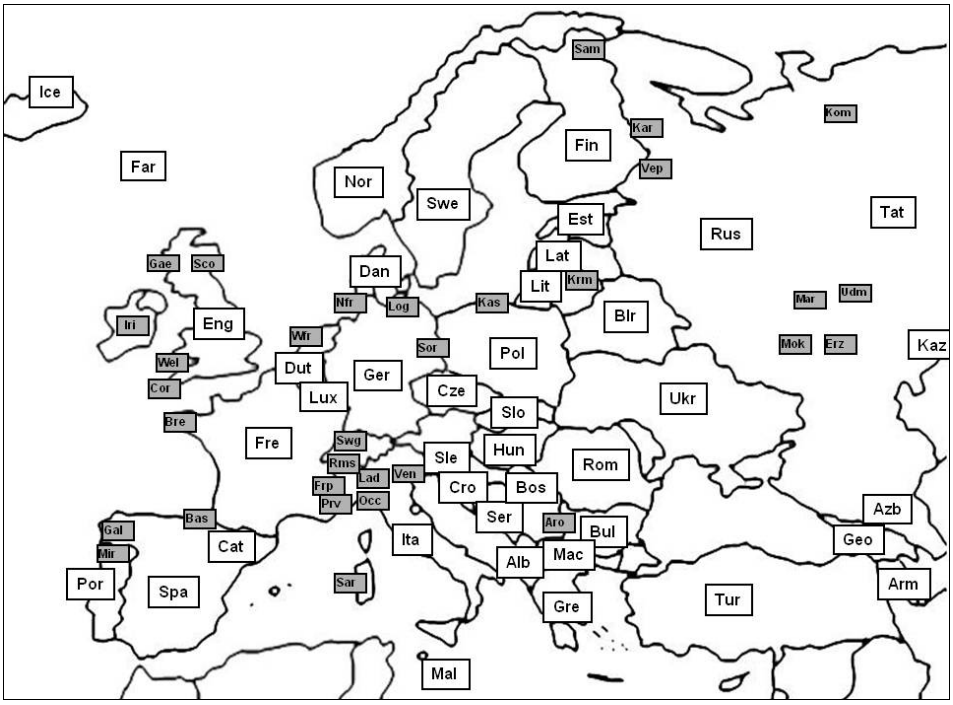
(iii) Turkic languages in Europe:
Our project includes five **Common Turkic** languages spoken in Europe: Two of them belong to the northwestern branch: the highly endangered or nearly extinct *Karaim* and *Tatar* which is under pressure from Russian and must be seen as "borderline case". *Turkish*, *Kazakh* and *Azerbaijani* are major languages. They belong to a southwestern subgroup of the Turkic languages.

(iv) Caucasian languages in Europe: Off the numerous Caucasian languages only *Georgian*, a language of the **Kartvelian** branch, takes part in the project. It is a major language.

(v) Semitic languages in Europe: *Maltese* belongs to the Semitic subgroup of the Afro-Asiatic languages; it is regarded as major language.

(vi) Basque: The isolated *Basque* is a lesser-used and partly endangered language.

Map 1 shows 73 European languages that are included in the project. Yiddish cannot be depicted on the map due to its ubiquity across the regions. The lesser-used languages are marked on gray rectangles, smaller than those of the major languages.



Map 1: Languages of the project "Widespread idioms in Europe and beyond"

3. Problems with data collection of lesser-used languages

As outlined in Section 1, our project is aimed at creating a comprehensive "Lexicon" of factually existing widespread idioms, including as many European languages as possible. In order to ascertain widespread idioms we have to rely on the assistance of many native speakers, linguistic groups and institutions. An important step was to build up a network of competent participants for many languages.

In the initial phase the "European Society of Phraseology" offered a good environment to realize the project: E-mails with information about the project and call for collaboration have been sent to members of this Society, and the first questionnaires were specially designed for this purpose. Following this, various individuals have helped with mediating experts of languages less accessible to idiom research. The questionnaires are also available on our website and several people who came across the homepage by chance have offered their cooperation. The questionnaires were modified many times and sent in several stages to experts of many languages, asking to answer questions based on both their own competence with regard to idioms and discussions within their circle of colleagues and families.

The collection of idioms of the major languages and the "borderline cases" Luxembourgish, Yiddish, Catalan and Galician does not require any further commentary. The

project has been supported by a number of competent colleagues and institutions. The questionnaires were filled in carefully, and often provided with additional comments on individual idioms. Several respondents have taken a lot of trouble to verify their information via investigations on the Internet, text-corpus analyses and dictionaries.

The minor and minority languages, however, will need some explication in terms of problems connected with collecting the data. We are dealing with two groups: On the one hand, there are lesser-used languages for which our collection is supported by an institution and extensive digital corpora are available. On the other hand, there are smaller languages where it is not possible to refer to phraseological dictionaries, databases or texts on the Internet, and we must content ourselves with the answers of individual respondents.

With regard to the first group Sorbian should be mentioned in the first place. Sorbian is well represented in our project, thanks to the reliable cooperation of the Sorbian Institute in Bautzen and Cottbus. The *Upper Sorbian* phraseology is present in a database that is continuously being expanded and available online on the Internet.[4] *Lower Sorbian* idioms are accessible online through the German-Lower Sorbian Dictionary (Deutsch-niedersorbisches Wörterbuch, DNW).[5] The situation is similar with *West Frisian*. The Fryske Akademy in Leeuwarden, the Netherlands with their large digital corpora supports our project. Recently, databases are also available for *North Frisian*, especially for the varieties of Amrum and Foehr, provided by the Nordfriisk Instituut in Bredstedt, Germany. For *Romansh*, the situation is also favorable. Our project is supported by the Lia Rumantscha in Chur, Switzerland, which can rely on a rich linguistic documentation of Romansh varieties. We got support from institutions and their competent researchers also for the lesser-used languages *Breton*, *Ladin*, *Mirandese* and *Kashubian.*

The data of the other lesser-used languages were collected by individual persons. Most of them are competent native speakers but some problems became obvious, especially in the realm of the highly endangered languages. Several respondents were not able to answer the questions about idioms themselves but had to ask their parents or grandparents for help. This was the case with a number of the smaller Romance and Northeast European languages where fluent speakers are aging, a major language is predominant, and transmission of the languages to the younger generations is not guaranteed. We have experienced this fact in our attempts to find competent informants. Various students from the republics of Komi, Udmurtia, Mari, and Mordvinia who we asked for cooperation did not feel competent enough to fill in the questionnaires or simply did not speak the "native" languages at all. In these cases we must consider the data given by their parents and grandparents as valuable sources.

Data of the following lesser-used languages were collected by individuals: *Scots*, *Low German*, *Swiss German* (Germanic); *Irish*, *Scottish Gaelic*, *Cornish*, *Welsh* (Celtic); *Provençal*, *Occitan*, *Francoprovençal*, *Venetian*, *Sardinian*, *Aromanian* (Romance); *Karelian*, *Veps*, *Udmurt*, *Komi-Zyrian*, *Mari*, *Moksha Mordvin*, *Erzya Mordvin*, *Inari Saam*i (Finno-Ugric); *Karaim*, *Tatar* (Turkic) as well as *Basque*.

When analyzing some widespread idioms we have to keep in mind this difference in collecting the data. The lack of idiom equivalents in the second group of lesser-used European languages means that no equivalents were known to our respondents and in

[4] Cf. WÖLKE (2008) and the Internet link.
[5] Cf. BARTELS (2013) and the Internet link.

most cases that there are no equivalents in these languages. Yet it cannot be excluded that an individual speaker would know an equivalent idiom if we asked a larger number of speakers.

4. Widespread idioms and the lesser-used languages: some examples

4.1. TO CUT THE GORDIAN KNOT

In the following we want to look at four examples in detail which can be considered representative of a larger number of widespread idioms. Let us look at the data of the idioms in full length, as they were reported by our respondents. The question was: "Is there an equivalent in your native language?" We chose idiom (1) as the first one because it belongs to an elevated, educated stylistic level and may illustrate the role of the smaller languages.

(1) English *to cut the Gordian knot* 'to solve an intractable problem in an amazingly simple and effective way; to make a quick and courageous decision; to resolve a difficult situation by force'

The idiom derived from reports on the deeds of Alexander the Great. They are about the legendary king Gordius of Phrygia, who dedicated his chariot to Zeus, binding it to a pole with a knot of great complexity that nobody could untie. When Alexander the Great passed through Phrygia in 333 BC, he heard about an oracle foretelling that whoever succeeded in untying the knot would become to rule over all of Asia. As the story tells, Alexander came and severed the knot into two with one stroke of his sword. This incident is reported by several authors. A summarizing phrase, *nodum solvere* "to undo the knot", had already become proverbial in antiquity and it has entered European vernacular languages.[6]

We will present the idioms according to the language families (cf. section 2) and emphasize the non-occurrence in the lesser-used languages. The result becomes visible in map 2 which shows gaps in various smaller languages.

(i) Indo-European languages in Europe

(a) Germanic languages:
No equivalents have been reported for four of the lesser-used Germanic languages (Scots, North Frisian, Low German, Swiss German) and Luxembourgish. For Yiddish the noun phrase *gordisher knup* 'gordian knot' was given. Only West Frisian possesses an equivalent: *de knoop trochhakje/trochhouwe/trochslaan* 'to chop/cut the knot'.
In contrast, the idiom exists in eight Germanic major languages:
Icelandic *að höggva á hnútinn* 'to chop the knot' / *(e-ð er) gordíonshnútur* '(sth. is) a Gordian-knot'
Faroese *høgga á (tann gordiska) knútin* 'to chop at (the gordian) knot'
Norwegian *hugge over den/en gordiske knute* 'to chop over the/a gordian knot'
Swedish *hugga av den gordiska knuten* 'to chop the gordian knot'

[6] Erasmus included the phrase in his "Adagia" (I i 6; Coll.W. 31, 54f). In his drama "King Henry V" (1, 1, 47; 1599), Shakespeare wrote "The Gordian knot of it he will unloose, Familiar as his garter". Cf. ALLEN 2006: 324 and 421; POHLKE 2008: 72.

Danish *løse/overhugge den gordiske knude* 'to untie/chop the gordian knot'
English *to cut the Gordian knot*
Dutch *de (Gordiaanse) knoop doorhakken* 'to chop the (Gordian) knot'
German *den (gordischen) Knoten durchhauen/durchtrennen* 'to chop/sever the (gordian) knot'

(b) Celtic Languages:
Four of the five lesser-used Celtic languages do not use an equivalent. For Cornish, an extinct but partly revitalized language, the form *treghy an colm fast* 'to cut the firm knot' was reported.

(c) Romance Languages:
According to our respondents, none of the nine lesser-used Romance languages possess an equivalent: neither Provençal, Occitan, Ladin, Romansh, Venetian, Sardinian, Francoprovençal, Mirandese nor Aromanian. For Catalan the noun phrase *nus gordià* 'gordian knot' was reported and for Galician the verb idiom *cortar o nó gordiano* 'to cut the gordian knot'.
All five Romance major languages use an equivalent:
French *trancher le nœud gordien* 'to cut the gordian knot'
Italian *tagliare/sciogliere un nodo gordiano* 'to cut/untie a gordian knot'
Spanish *cortar el nudo gordiano* 'to cut the gordian knot'
Portuguese *cortar o nó górdio* 'to cut the gordian knot'
Romanian *a tăia/a reteza/a desface/a dezlega nodul gordian* 'to cut/to loosen/to untie the Gordian knot'

(d) Baltic Languages: Equivalents exist in both languages.
Latvian *pārcirst Gordija mezglu* 'to cut off the Gordian knot'
Lithuanian *Gordijaus mazgą perkirsti* 'to cut off the Gordian knot'

(e) Slavonic Languages:
No equivalent was recorded for the declining Kashubian while both Sorbian varieties have an equivalent: Upper Sorbian *gordiski suk přerubnyć* and Lower Sorbian *gordiski suk rozbiś*, both meaning literally 'to cut the gordian knot'.
All twelve Slavonic major languages of the project are represented:
Russian *перерубить/разрубить Гордиев узел* 'to cut the Gordian knot'
Belorussian *перасячы гордзіеў вузел* 'to cut the gordian knot'
Ukrainian *розрубати Гордіїв вузол* 'to cut the Gordian knot'
Czech *rozetnout gordický uzel* 'to cut the gordian knot'
Slovak *preťať gordický uzol* 'to cut the gordian knot'
Polish *przeciąć węzeł gordyjski* 'to cut the gordian knot'
Slovene *presekati gordijski vozel* 'to cut the gordian knot'
Croatian *presjeći gordijski čvor* 'to cut the gordian knot'
Bosnian *presjeći gordijski čvor* 'to cut the gordian knot'
Serbian *пресећи Гордијев чвор* 'to cut the Gordian knot'
Macedonian *го пресече Гордиевиот јазол* 'to cut the Gordian knot'
Bulgarian *разсичам гордиевия възел* 'to cut off the gordian knot'

(f) Albanian: *zgjidh nyjën gordiane* 'to solve the Gordian knot'

(g) Greek: *λύνω το Γόρδιο δεσμό* ‘to solve the Gordian knot’

(h) Armenian: *gordjan hangujc* ‘the Gordian knot’

(ii) Finno-Ugric Languages in Europe

(a) Ugric Languages:
Hungarian has an equivalent: *kettévágja/megoldja a gordiuszi csomót* ‘sb. cuts in half/solves the gordian knot’

(b) North-Finnic Languages:
The lesser-used Karelian has no idiom while an eqivalent was given for the highly endangered Veps: *pästta sol’mespäi* ‘to extricate from the knot’.
Both North-Finnic major languages use equivalents:
Finnish *selvittää Gordonin solmu* ‘to solve the Gordian knot’
Estonian *Gordioni sõlme lahti raiuma* ‘to cut off the Gordian knot’

(c) Permic Languages:
No equivalent was given for Udmurt while Komi uses the idiom *джаггӧрӧд ӧрӧдны* ‘to break the tightened knot’.

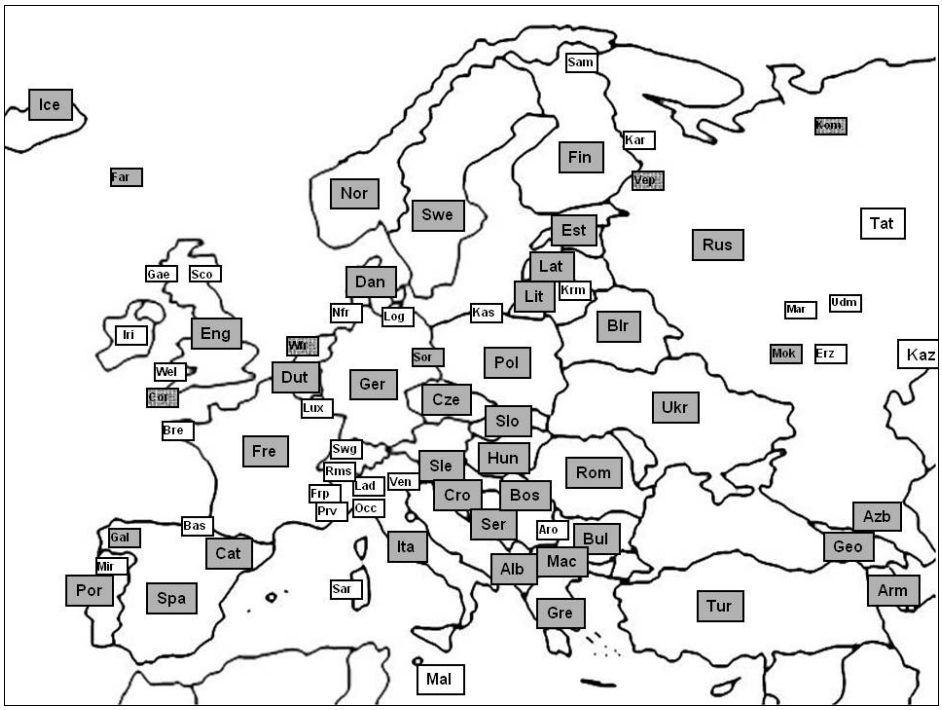
(d) Volgaic Languages:
No equivalents were given for Mari and Erzya Mordvin while Moksha-Mordvin is represented by the idiom *Гордиень сюлмоть орадомс* ‘to cut the Gordian knot’.

(e) Saamic Languages: No equivalents.

(iii) Turkic Languages in Europe
The highly endangered Karaim as well as Tatar and Kazakh are not represented in our data. Equivalents were recorded for two Turkic standard languages:
Turkish *Gordion düğümünü çözmek* ‘to untie the Gordian knot’
Azerbaijani *iskəndər düyünü* ‘Alexander’s knot’

(iv) Caucasian languages in Europe
Georgian has an idiom: *gordiseuli/gordiusis k’vandzis gakhsna* ‘to sever the gordian knot/Gordius’ knot’

(v) Maltese and (vi) Basque: no equivalents.
The material sent in by our informants reveals that equivalents of the idiom are fairly widespread across the European languages, ranging from Icelandic to Armenian or from Portuguese to Moksha Mordvin. Parallels exist even in non-European languages, cf. Korean *gorudiosui maedup* ‘the gordian knot’ and Chinese *kuàidāo zhǎn luànmá* ‘sharp knife – to cut knot’. The Chinese idiom and the Azerbaijani idiom above (‘Alexander’s knot’) show that not all idioms are borrowings or loan translations but can go back independently of each other to the same source, here to fragments of world knowledge about Alexander’s story. Apart from these examples the idioms are quite consistent in their lexical and morpho-syntactic structure. There are only small variants concerning the verb (CUT, CHOP, SEVER, UNTIE, SOLVE, etc.).



Map 2: Equivalents of *to cut the Gordian knot* in European languages

What is more striking, however, is the fact that in some languages, GORDIAN is an optional addition. This results in an "educated" high-register variant and a colloquial variant side by side. Thus, the English *to cut the knot* or German *den Knoten durchhauen* can be used with an allusion to cutting the Gordian knot (cf. Latin *nodum solvere*) or independently of it. Such idioms could even be developed independently of our idiom, since KNOT can be used as a metaphor for something complicated (cf. idioms like *to tie yourself (up) in knots* 'to become very confused or worried when you are trying to make a decision or solve a problem'). Our respondents of West Frisian, Cornish, Veps and Komi reported idioms without the element GORDIAN. Probably, these idioms came into being independently of the present widespread idiom.

Let us summarize: Our idiom is a good example to illustrate the role of the lesser-used languages. We encountered idiom equivalents in 36 major languages, in the "borderline cases" Yiddish, Catalan and Galician and in six lesser-used languages. Four of them do not refer to GORDIAN and, therefore, probably must be separated. The remaining lesser-used languages are Moksha Mordvin (for which we have no explanation) and the Sorbian varieties for which there is access to digital corpora.[7] Map 2 is nearly congruent with map 1, but the other way round: rectangles that mark the smaller languages remain white. The idiom belongs to an erudite language due to its classical connotations. Therefore, it did not find its way into lesser-used languages without long literary traditions. This example is not an isolated case, but has various parallels in idioms that have been studied so far.

[7] The Dictionary of Lower Sorbian documents the sentence *Jomu jo se napšawdu raźiło gordiski suk rozbiś* "he actually managed to cut the gordian knot".

4.2. A BLACK DAY

Several widespread idioms do not belong to a high stylistic register but are rather "unobtrusive" with regard to their origin. However, the projection of the data on the map of Europe produces a similar picture as map 2. Let us discuss another widespread idiom in view of the lesser-used languages involved. The idiom (2) is rooted in superstitious ideas of the Romans, who classified the week into black and white days, i.e. in unfavorable and happy days, and marked certain disastrous days in the calendar. The historian Livy, the philosopher Seneca and other Roman writers wrote in detail about such days in the Roman history.[8] The color BLACK has the general symbolic function of 'bad', which contributes to the easy understanding of the idiom's figurative meaning.

(2) *a black day* 'a day full of failures or disasters'

We will present the data in a simplified schematic form in the order of firstly "lesser-used languages", secondly "borderline cases" (only for Germanic, Romance and Turkic) and thirdly "major languages". We choose a shortened arrangement for the Finno-Ugric languages. The following idioms literally mean '(a) black day' unless noted differently.

(i) Indo-European

(a) Germanic:
Lesser-used: No equivalents for Scots and North Frisian. West Frisian *in swarte dei*, Low German *en schwatte Dag* and Swiss German *e schwarze Taag.*
Borderline cases: Luxembourgish *e schwaarzen Dag*, Yiddish *a shvartsn tog.*
All eight major languages: Icelandic *svartur dagur*, Faroese *svartur dagur*, Norwegian *en svart dag*, Swedish *en svart dag*, Danish *en sort dag*, English *a black day*, Dutch *een zwarte dag*, German *ein schwarzer Tag.*

(b) Celtic:
Lesser-used: No equivalents for Gaelic and Cornish. Irish *lá dubh*, Welsh *diwrnod du*, Breton *un devezh du.*

(c) Romance:
Lesser-used: No equivalents for Ladin, Francoprovençal, Venetian, Mirandese, and Aromanian. Provençal *una journa neira*, Occitan *unë zhournà nhêrë*, Romansh *in ner gi*, Sardinian *un die négra.*
Borderline cases: Catalan *un dia negre*, Galician *un día negro.*
All five major languages: French *une journée noire*, Italian *un giorno nero / una giornata nera*, Spanish *un día negro*, Portuguese *um dia negro*, Romanian *o zi neagră*.

(d) Baltic: Latvian *melnā diena*, Lithuanian *juoda diena*.

(e) Slavonic:
Lesser-used: No equivalent for Kashubian. Upper Sorbian *čorny dźeń*, Lower Sorbian *carny źeń.*
All twelve major languages: Russian *чёрный день*, Belorussian *чорны дзень*, Ukrainian *чорний день*, Czech *černý den*, Slovak *čierny deň*, Polish *czarny dzień*, Slovene *črn dan*, Croatian *crni dan*, Bosnian *crni dan*, Serbian *црни дан*, Macedonian *црн ден*, Bulgarian *черен ден.*

[8] Cf. POHLKE/POHLKE 2006: 142 for more detail.

(f) Albanian: *një ditë e zezë*

(g) Greek: *μαύρη μέρα*

(h) Armenian: *sev or*

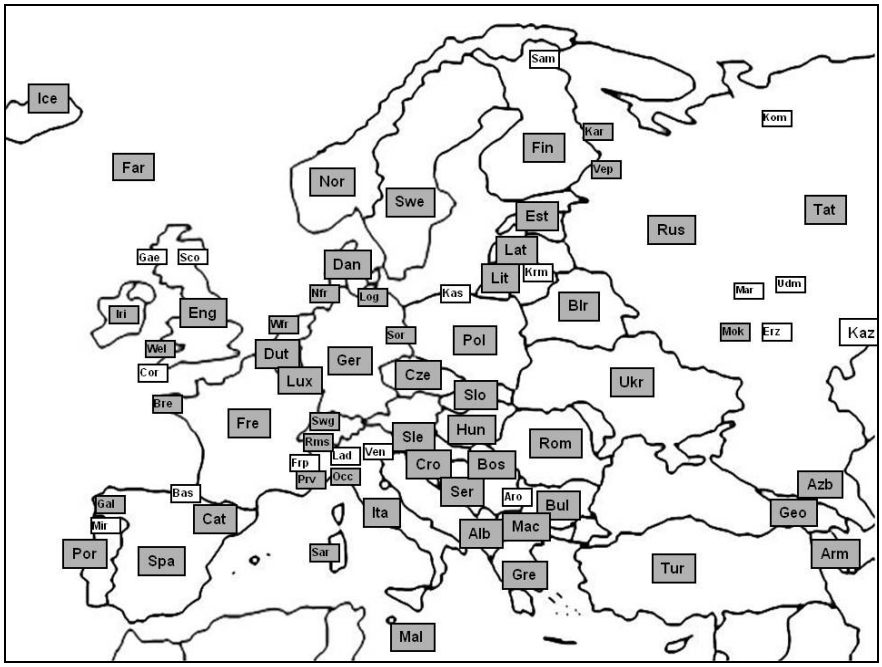
(ii) Finno-Ugric
Lesser used: No equivalents for Komi, Udmurt, Mari, Erzya Mordvin and Inari Saami. Karelian *mustat päivät* "black days", Veps *mustad päivät* "black days", Moksha Mordvin *равжа шш.*
All three major languages: Hungarian *fekete nap*, Finnish *musta päivä*, Estonian *must päev.*

(iii) Turkic
Lesser used: No equivalent for Karaim.
Tatar: no equivalent.
Major languages: Turkish *kara gün*, Azerbaijani *qara gün.* No equivalent for Kazakh.

(iv) Caucasian: Georgian *shavi dghe*.

(v) Semitic: Maltese *ġurnata sewda.*

(vi) Basque: No equivalent.
To sum up: Equivalents are spread across at least 56 European languages in a very consistent lexical structure and figurative meaning. All major languages included in our project except for Kazakh possess the idiom. However, various smaller languages do not use any equivalent, possibly because tacit knowledge of the origin of the idiom points to a high register of language for which there is no place in the lesser-used or declining languages.



Map 3: Equivalents of *a black day* in European languages

4.3. TO BE SITTING ON A POWDER KEG

Let us discuss another widespread idiom which comes from the source domain WARFARE, (3). Powder kegs were small barrels used to store and transport large quantities of gunpowder. These kegs had to be handled with care because their contents could go off at any moment. The image of sitting on such a highly explosive barrel serves as a means of comparison with a dangerous situation that can lead to a violent upheaval.

(3) *to be sitting on a powder keg* 'to be in a very dangerous situation in which something could suddenly go seriously wrong at any time'

We will present the data in the shortened form as for example (2). The Germanic idioms below mean literally 'to sit on a powder-barrel' unless noted differently. Most of the Romance idioms below translate literally as 'to be on a powder-barrel'; this meaning has been omitted. The literal meaning of the Slavonic idioms 'to sit on a barrel of gunpowder' has also been omitted.

(i) Indo-European

(a) Germanic:
Lesser used: No equivalents for North Frisian and Low German. Scots *sittin on a pouther keg* 'sitting on a powder keg', West Frisian *krûdfet* 'powder barrel', Swiss German *uf eme Pulverfass sitze*.
Yiddish and Luxembourgish: no equivalents.
Major languages: Icelandic *sitja á púðurtunnu*, Norwegan (Bokmål) *sitte på en kruttetønne* / (Nynorsk) *sitje på ei kruttønne*, Swedish *sitta på en krutdurk*, Danish *sidde på en krudttønde*, English *to be sitting on a powder keg*, Dutch *op een kruitvat zitten*, German *auf einem Pulverfass sitzen*. No equivalent for Faroese.

(b) Celtic: No equivalents.

(c) Romance:
Lesser used: No equivalents for Occitan, Ladin, Francoprovençal, Venetian, Sardinian, Mirandese and Aromanian. Provençal *èstre sus uno poudriero.*
Catalan *seure sobre un barril de pólvora*, Galician *ser unha polvoreira/un polvorín* "to be a powder-barrel".
Major languages: French *être sur une poudrière*, Italian *star seduti su una polveriera* 'to be sitting on a powder-barrel', Spanish *estar sentado en un polvorín* 'to be sitting on a powder-barrel', Portuguese *estar em cima dum barril de pólvora* 'to be on top of a barrel of powder' / *estar sobre um barril de pólvora* 'to be on a barrel of powder'. No equivalent for Romanian.

(d) Baltic: Latvian *sēdēt (kā) uz pulvera mucas* '(like) to sit on a/the barrel of powder', Lithuanian *sėdėti ant parako statinės* 'to be sitting on a barrel of powder'.

(e) Slavonic:
Lesser used: No equivalents for Kashubian. Upper Sorbian *pólwrowy sud* 'powder keg', Lower Sorbian *sejźeś na pulbjerowym suźe.*
Major languages: Russian *сидеть на пороховой бочке* 'to sit on a powder keg', Ukrainian *сидіти на пороховій бочці/на пороховому погребі* 'to sit on a powder keg/on a gunpowder magazine', Czech *sedět na sudu s prachem* 'to sit on a barrel of powder',

Slovak *sedieť na sude s pušným prachom*, Polish *siedzieć na beczce prochu*, Slovene *sedeti na sodu smodnika*, Croatian *sjediti na bačvi baruta*, Bosnian *sjediti na buretu baruta*, Serbian *седети на буретy барута*, Macedonian *седи на буре барут*. No equivalents for Belorussian and Bulgarian.

(f) Albanian: *te ulesh mbi fuci baruti* 'to sit on a barrel of gunpowder'

(g) Greek: *είναι/θυμίζει πυριδιταποθήκη* 'to be/remember of a powder magazine'

(h) Armenian: *varodí takár* 'a powder keg'

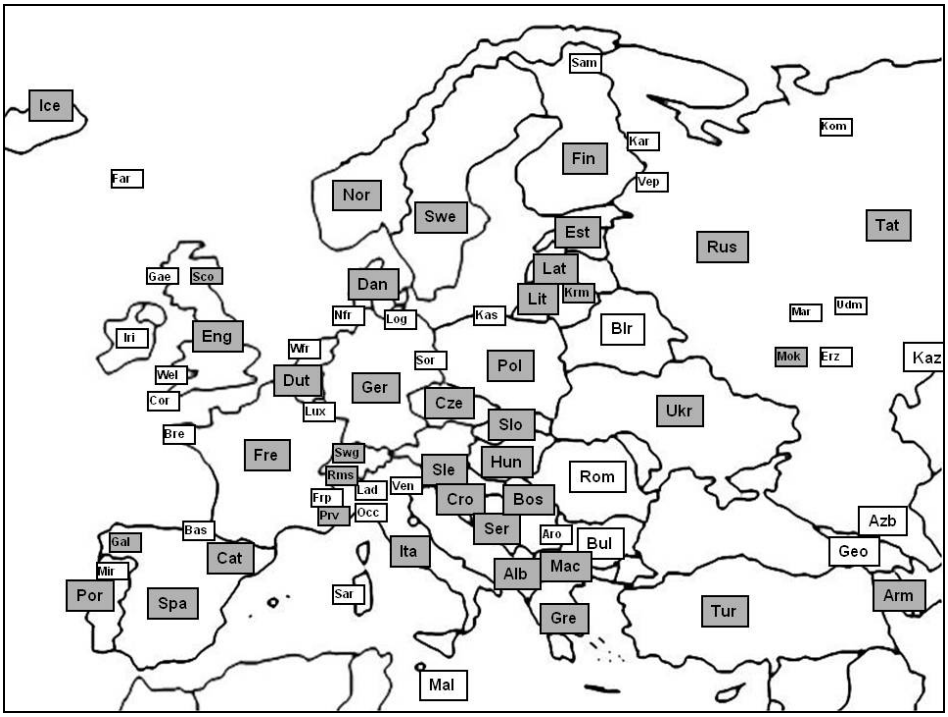
(ii) Finno-Ugric
Lesser used: No equivalents for Karelian, Veps, Komi, Udmurt, Mari, Erzya Mordvin and Inari Saami. Moksha Mordvin *порохонь мархта боцькя лангса ащемс* 'to be on a barrel of gunpowder'.
All three major languages: Hungarian *puskaporos hordón ül* 'sb. sits on a powder-barrel', Finnish *tuntea istuvansa ruutitynnyrillä* 'to feel like sitting on a powder-barrel' and Estonian *püssirohutünni otsas istuma* 'to be sitting on the end of a powder-barrel' / *nagu püssirohutünni otsas elama* 'like to live on a powder-barrel'.

(iii) Turkic
Lesser-used: Karaim *barot čapčahy üstiuńdia olturma* 'to sit on a powder barrel'.
Tatar *дары мичкэсе остендэ утыру* 'to be sitting on a powder keg'.
Major languages: Turkish *barut fiçısının üstünde oturmak* 'to sit on a powder barrel'.
No equivalents for Kazakh and Azerbaijani.

(iv) Caucasian (Georgian), **(v) Semitic (**Maltese) and **(vi) Basque:** No equivalents.
Our research shows that equivalents of the idiom occur in a large number of European languages with a very consistent form and meaning. Some non-European languages use near-equivalent idioms as well. Examples include Farsi *rooyeh boshkeh baroot neshatan* 'sitting on a powder barrel', Chinese *zuòzài huǒyàotǒng shang* 'to sit on a powder barrel', Vietnamese *ngồi trên thùng thuốc nổ* 'to sit on a powder barrel' (less common) and, more common, *ngồi trên đống lửa* 'to sit on a fire pile' and Japanese *kayakuko no ue ni koshikakete iru* 'to be sitting on a powder magazine'. Compare also Korean *hwayak-eul jigo bulro deuleo-gada* 'to carry gunpowder in the back (and) go into the fire'.

The noun POWDER-BARREL can be used metaphorically for 'an imminent danger, an explosive situation', even in languages where the idiom itself does not exist (cf. the West Frisian, Galician, Upper Sorbian, Greek and Armenian expressions above). These metaphors often occur in the context of the danger of a war, of territories affected by an outburst of violence, where the literal meaning of GUNPOWDER (as used in war) is simultaneously acknowledged. Morpho-syntactic additions like 'social powder keg' or 'demographic powder keg' are common in the major languages.

This example shows that nothing is predictable. Although there are clear gaps in the lesser-used languages equivalents have been reported for Scots, Swiss German, Provençal, Moksha Mordvin and the highly endangered Karaim, while six major languages (Faroese, Romanian, Belorussian, Bulgarian, Kazakh and Azerbaijani) do not use the idiom.



Map 4: Equivalents of *to be sitting on a powder keg* in European languages

4.4. TO PLAY A ROLE (FOR SOMEONE/SOMETHING)

The last idiom to be discussed here is derived from the THEATER and goes back to a detail of acting. Starting from the end of the 16th century, the text of actors used to be written on paper rolls, later called ROLEs. During rehearsals, the piece of text that the actors were actually using would be visible, while the rest would be rolled up. Through this, the concept ROLE evolved into the more general sense of 'personal appearance and work, individual performance (as if in a play)' and it is this meaning that appears in a number of idioms, cf. (4).

(4) *to play a role (for someone/something)* 'to be of some importance (for sb./sth.); to be involved (in sth.), to have a specific effect (on sb./sth.)'

The following idioms translate as 'to play a role' unless noted differently.

(i) Indo-European

(a) Germanic:
Lesser used: No equivalent for Scots. North Frisian *dat spält en rol* 'that plays a role', West Frisian *in rol/rolle spylje*, Low German *en Rull spelen*, Swiss German *e Rolle spile.*
Luxembourgish *eng Roll spillen*, Yiddish *shpiln a role.*

All eight major languages: Icelandic *leika hlutverk*, Faroese *tað spælir en rollu* 'that plays a role', Norwegian *spille en rolle*, Swedish *spela en roll*, Danish *spille en rolle*, English *to play a role*, Dutch *een rol spelen*, German *eine Rolle spielen.*

(b) Celtic:
Lesser used: No equivalents for Irish, Gaelic and Cornish. Welsh *chwarae rôl*, Breton *c'hoari ur roll*.

(c) Romance:
Lesser used: No equivalents for Ladin, Francoprovençal, Venetian and Sardinian. Provençal *jouar un rôl*, Occitan *aiquë un rolë* 'to have a role', Romansh *giugar ina rolla*, Mirandese *tener un papel* 'to have a role', Aromanian *lj-si mãcã zborlu.*
Catalan *jugar un paper*, Galician *xogar un papel.*
All five major languages: French *jouer un rôle*, Italian *avere un ruolo* 'to have a role', Spanish *jugar/desempeñar un papel*, Portuguese *ter/desempenhar um papel* 'to have/play a role', Romanian *a juca un rol.*

(d) Baltic: Latvian *spēlēt lomu*, Lithuanian *vaidinti vaidmenį.*

(e) Slavonic:
Lesser used: Kashubian *grac rolą*, Upper Sorbian *rólu hrać*, Lower Sorbian *žednu rolu njegraś* 'to play no role'.
All twelve major languages: Russian *играть роль*, Belorussian *адыгрываць роль*, Ukrainian *играть роль*, Czech *hrát roli*, Slovak *hrať úlohu*, Polish *odgrywać rolę*, Slovene *igrati vlogo*, Croatian *igrati ulogu*, Bosnian *igrati ulogu*, Serbian *играти улогу*, Macedonian *igra uloga*, Bulgarian *играя роля.*

(f) Albanian: *luan rol*

(g) Greek: *κάτι παίζει ρόλο* 'sth. plays a role'

(h) Armenian: *der khaghál*

(ii) Finno-Ugric
Lesser used: No equivalents for Veps, Komi, Mari, Erzya Mordvin and Inari Saami. Karelian *hänäl on ruoliloi* 's/he has a role', Udmurt *нокыче но уг пот* 'to play no role' / *(аслэсьтыз) рользэ шудыны* 'to play (one's) role', Moksha Mordvin: *роль налхкомс.*
All three major languages: Hungarian *szerepet játszik* 'sth. plays a role', Finnish *näytellä/esittää jnklaista roolia* 'to play/present some kind of role', Estonian *mingit rolli mängima* 'to play no role'.

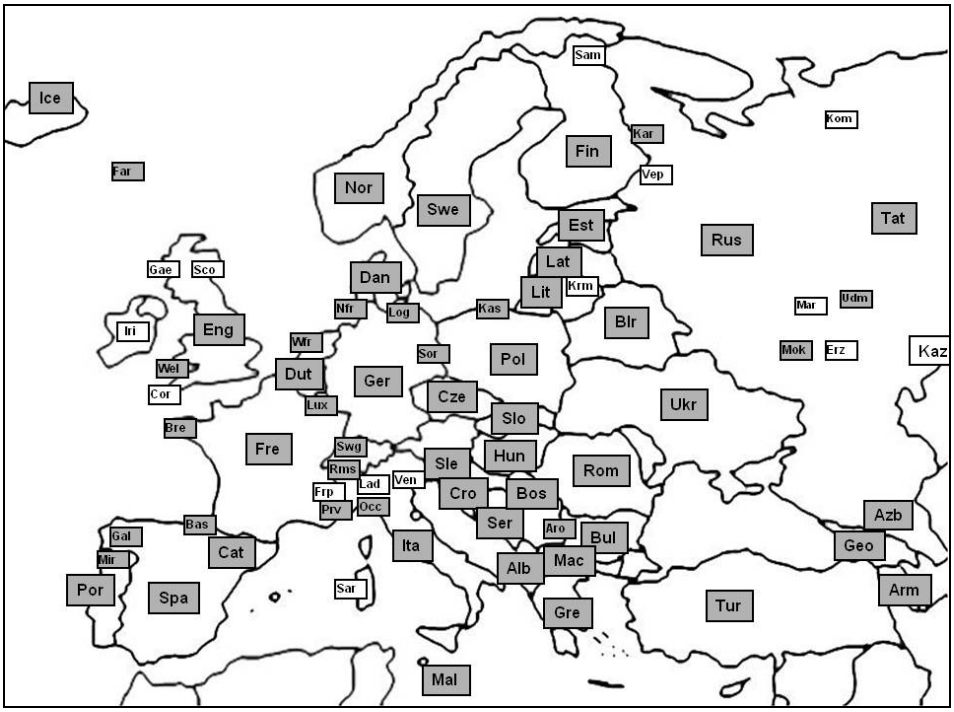
(iii) Turkic
Lesser used: No equivalent for Karaim.
Tatar *роль уйнау*
Major languages: Turkish *rol oynamak*, Azerbaijani *rol oynamaq*. No equivalent for Kazakh.

(iv) Caucasian: Georgian *rols tamashobs*

(v) Semitic: Maltese *jilgħab rwol*

(vi) Basque: *en papera egin*
The idiom is very familiar to speakers of the major languages. It is stylistically neutral and adapts to all sorts of contexts. Therefore, it is unobtrusive, even to the point of not being perceived as an idiom at all. It occurs in different variants, and views diverge as to which one may be the most frequent one. Some respondents assume that the negative version 'to play no role' is more common (e.g. Danish *ngt. spiller ingen rolle* 'sth. does not play any role' or Croatian *ne igrati nikakvu ulogu* 'not to play any role', compare also the Lower Sorbian idiom above). In most cases, however, the idiom is used with an adjective, such as 'to play a great/an important role', 'to play the most significant role', 'to play a subordinate role', etc.

To sum up: Equivalents of the idiom were reported for all major languages (except for Kazakh), all "borderline cases" and 17 lesser used languages represented in our project. The idiom is also encountered in several non-European languages.[9] It is thus one of the most popular widespread idioms. We have no clear explanation why some of the smaller languages do not possess an equivalent, possibly because the source concept THEATER/SPECTACLE still points to a high register of language or, more generally, because of their status as declining languages.



Map 5: Equivalents of *to play a role* in European languages

[9] Full equivalents in non-European languages include, for example: Chinese *bànyǎn yīgè jiǎosè* 'a role to play', Mongolian *үүрэгтэй байх* 'to play a role', Korean *yeog-eul ha-da* 'to make a role', Japanese *juuyou na yakuwari wo enzuru* 'to play an important role'.

5. Outlook

The study on "Widespread Idioms" shows that the languages spoken at the geographical periphery of Europe (from Icelandic and Norwegian to Turkish, Georgian and Armenian) are not marginal figures for the most part but central to a "Lexicon of Common Figurative Units" while most lesser-used languages, even located in the middle of Europe (e.g. Kashubian, several Alpine Romance languages, Aromanian, etc.) are clearly outside. This provides us with new insights into the far-reaching uniformities of idioms of the European languages. The minority varieties are not integrated into the European uniformity to the same degree as is the case with the major literary languages. Various widespread idioms that are well-developed in the major languages are hardly shared by the lesser-used languages. This may be due to an elevated style of such idioms (cf. the "educated" idiom *to cut the Gordian knot* originating from an event in ancient history and legends developed from it) but also to tacit knowledge of their once high-registered origin (*a black day* going back to Roman superstition), or to other special source domains such as WARFARE (*to be sitting on a powder keg*) and THEATER (*to play a role*), deputizing here for a number of widespread idioms with a similar distribution across the languages of Europe.

There are several attempts at explaining the causes of the spreading of idioms across many languages. One of them is the assumption of polygenetic origin of several idioms. This may apply to a very small number of widespread idioms which are based on common human experience or observation of nature and environment.[10] Another explanation often put forward is the increasing influence of English. This modern development of the 20th century, however, can be ruled out for the questions we are asking here since our widespread idioms belong to much older layers. An alternative explanation may be the well-known contact linguistic model of lexical transfer[11] which emphasizes constant loan translations and borrowings among languages that are in close geographic and cultural contact to one another, especially in situations of bilingualism. However, this model alone cannot explain the Europe-wide dissemination of an idiom. We did not detect a "center", a dominant source language from which the idioms would have found their way into all the languages under consideration.

However, the contact linguistic model has an explanatory potential for the existence of widespread idioms in the lesser-used languages, at least to some extent. It explains restricted regional borrowings, usually from major to minor languages due to bilingualism.[12] As VEISBERGS (2012: 65) points out, phraseological borrowings are distinct from lexical borrowings. "In lexical borrowing, the new lexeme introduced in the receptor language is alien and incomprehensible to native speakers, whereas in phraseological

---

[10] An example of independent development is (*at*) *day and night* meaning figuratively 'all the time, continually, perpetually'. An indication of polygenesis is the fact that equivalents are not only known in almost all of the lesser-used languages (cf. Upper Sorbian *wodnjo a w nocy* 'at day and at night') but also in numerous non-European languages (cf. PIIRAINEN 2012: 316–320).

[11] Since WEINREICH's (1953) seminal work, a wealth of studies on this subject have appeared. Good overviews of the different processes of borrowing and their stages can be found in FIELD 2002, PUDA 2010, among others, and especially for phraseological borrowing in VEISBERGS 2012.

[12] In a bilingual situation, also the reverse is possible. For instance, idioms of an old local dialect can – often unnoticed – enter into the (regional colloquial) major languages. See PIIRAINEN (2008) for German examples.

borrowing (mostly loan translations) a PU[13] is often motivated and comprehensible" (ibd.). Most speakers of the European lesser-used languages are currently bilingual. This explains that idioms (as motivated linguistic units) are easily acquired from the overarching major languages by the smaller ones. With regard to our project, idioms may be borrowed from Lithuanian to Karaim, from English to several Celtic languages, from French to Breton and Provençal, from Italian to Venetian and Sardinian, from Russian into various North-East European languages, from Polish to Kashubian, from German to Sorbian, and so on. However, we must not draw far-reaching conclusions on the basis of the linguistic data alone, which were collected synchronically by our surveys. Historical analyses, examining the idioms' development from the earliest instance onwards, would be needed to answer questions as to how a donor language can be identified, and how an idiom can be judged to be a borrowed one.

Through our research on widespread idioms, we are able to provide information on the question as to what extent the figurative lexicon of these lesser-used languages shares common features with that of the major literary languages. West Frisian and Upper Sorbian reveal the greatest consistency with the "Lexicon of Common Figurative Units", while the endangered languages Inari Saami and Komi as well as Mari – which stem from cultural traditions that are very different from those of most of the other European languages – are at the opposite end. The declining Cornish and other almost extinct varieties show a similar picture, due to other causes.

The search for similarities of the figurative lexicon across many languages, however, should not obscure the view for an even more important linguistic topic, i.e. the study of the figurative lexicon of those languages which are on the verge of extinction. It is expected that half of the world's languages will disappear forever within the present century. Figurative units such as idioms are among the most vulnerable elements of language; they begin to vanish in the first phase of a language becoming endangered.[14] The documentation and research on these "vanishing figurative units" is an even more urgent task for linguists in the future than is the study of widespread idioms in Europe and beyond.

Bibliography

ALLEN, ROBERT (2006): *Allen's Dictionary of English Phrases*. Los Angeles: Penguin Books.

BARTELS, HAUKE (2013): Zur Konzeption eines historisch-dokumentierenden Wortschatz-Informationssystems des Niedersorbischen. Pläne zur Behebung eines drängenden Forschungsdesiderats. In: KEMPGEN, S.; WINGENDER, M.; FRANZ, N.; JAKIŠA, M. (Hrsg.): *Deutsche Beiträge zum 15. Internationalen Slavistenkongress, Minsk 2013* (Die Welt der Slaven. Sammelbände. Sborniki; 50). München [etc.]: Otto Sagner, 37–46.

CRYSTAL, DAVID (2000): *Language Death*. Cambridge: Cambridge University Press.

[13] PU is the abbreviation of *phraseological unit*.

[14] Cf. footnote 3 and (IDSTRÖM/PIIRAINEN 2012) for more details. The excellent documentation of the Sorbian phraseology should encourage researchers of other lesser-used languages to regard the documentation of the figurative lexicon of these languages as an object of urgency.

ERASMUS COLL.W.: *Collected Works of Erasmus*, Vol. 31 Adages Ii1 to Iv100 (1982), translated by Margret Mann Phillips, annotated by R. A. B. Mynors. Toronto [etc.]: Toronto University Press.

FIELD, FREDRIC (2002): *Linguistic Borrowing in Bilingual Contexts.* With a foreword by Bernard Comrie. Amsterdam/Philadelphia: John Benjamins.

HARRISON, DAVID K. (2007): *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge*. New York [etc.]: Oxford University Press.

HARRISON, DAVID K. (2010): *The Last Speakers. The Quest to Save the World's Most Endangered Languages.* Washington DC: National Geography.

IDSTRÖM, ANNA; PIIRAINEN, ELISABETH (2012): Endangered Metaphors: Introduction. In: IDSTRÖM, A; PIIRAINEN, E. (eds.): *Endangered metaphors*. Amsterdam: John Benjamins, 15–19.

NETTLE, DANIEL; ROMAINE, SUZANNE (2000): *Vanishing Voices: The Extinction of the World's Languages*. Oxford: Oxford University Press.

PIIRAINEN, ELISABETH (2008): Phraseology from an areal linguistic perspective. In: MARÍA ÁLVAREZ DE LA GRANJA (ed.): *Fixed Expressions in Cross-Linguistic Perspectives. A Multilingual and Multidisciplinary Approach.* Hamburg: Kovač, 19–44.

PIIRAINEN, ELISABETH (2012): *Widespread Idioms in Europe and Beyond. Toward a Lexicon of Common Figurative Units*. New York: Peter Lang.

POHLKE, REINHARD (2008): *Das wissen nur die Götter. Deutsche Redensarten aus dem Griechischen.* Düsseldorf: Patmos Verlag.

POHLKE, REINHARD; POHLKE, ANNETTE (2006): *Alle Wege führen nach Rom. Deutsche Redensarten aus dem Lateinischen.* Düsseldorf: Patmos Verlag.

PUDA, ALEŠ (2010): *Zur Theorie der Lehnprägung im deutsch-tschechischen Sprachkontakt. Eine historisch-vergleichende Untersuchung im innerslavischen und europäischen Kontext.* Frankfurt/M.: Peter Lang.

SHERZER, JOEL; STOLZ, THOMAS (eds.) (2003): *Minor Languages. Approaches, Definitions, Controversies. Papers from the conference on 'Minor Languages: Coming to grips with a suitable definition'*, Bremen, June 2001. Bochum: Brockmeyer.

VEISBERGS, ANDREJS (2012): *Phraseological Borrowing.* Berlin: Logos Verlag.

WEINREICH, URIEL (1953): *Languages in Contact. Findings and Problems.* With a Preface by André Martinet. New York: Publications of the Linguistic Circle.

WÖLKE, SONJA (2008): Hornjoserbska frazeologija w interneće [Obersorbische Phraseologie im Internet]. In: *Lětopis, Zeitschrift für Sorbische Sprache, Geschichte und Kultur* 55/1, 36–47.

Internet:

http://www.serbski-institut.de/cms/de/50	Obersorbisches phraseologisches Wörterbuch
http://www.dolnoserbski.de/dnw/index.htm	Deutsch-Niedersorbisches Wörterbuch DNW
http:// www.widespread-idioms.uni-trier.de	Project Widespread Idioms in Europe and Beyond
http://www.unesco.org/culture/languages-atlas/	UNESCO Atlas of the World's Languages in Danger