

*Elisabeth Piirainen**

***Europeanism, internationalism or something else?
Proposal for a cross-linguistic and cross-cultural
research project on widespread idioms in Europe
and beyond***

Abstract

Terms like **Europeanism** or **internationalism** are used rather carelessly in phraseology research, and they are often applied to idioms accidentally observed to have similar lexical and semantic structures in a few languages. The central question as to what idioms actually constitute the so-called “phraseological uniformity of Europe” has never been studied systematically. The present paper describes my first steps of a research project aimed at creating an inventory of factually existing **widespread idioms**. The total set of idioms is subdivided into smaller groups on the basis of their cultural foundations, thus becoming comparable across a great variety of languages. Questionnaires sent to a large number of idiom research experts are producing first results: numerous idioms are spread across at least several dozens of languages. The problem remains that idioms are currently accessible for only a small portion of the approx. 140 European languages.

**1. By way of introduction: proverb research
and idiom research**

The present article is concerned with idioms that are widespread across various languages. Phraseology researchers appear to have some vague knowledge about a cross-linguistic uniformity of idioms in European languages, but the set of problems connected with this concept has never been methodically studied. Some researchers even use terms like **Europeanism** or **internationalism** for idioms that happen to be found “similar” – more or less accidentally – in a small number of languages.

* *Dr. Elisabeth Piirainen*
Dumte 32
D-48565 Steinfurt
piirainen@t-online.de

This paper reports on a recent attempt to look into this phenomenon in a more systematic way.

Since this article will concentrate on idioms proper, a short comment on the term **idiom** seems appropriate. In English, idiom is sometimes used as an umbrella term for all kinds of fixed expressions, cf. Čermák (2004), who speaks of “proverbs and other idioms”. However, this usage is quite misleading because it is ambiguous and contradictory to the usage of terms such as **Idiom** in German or **идиома (idioma)** in Russian. In this study, idiom will be understood in a narrower sense and reserved for multi-word units that are highly irregular, i.e. show a high degree of stability and idiomaticity. Idioms are seen as the central category of phraseology, distinguished from other categories such as proverbs and less idiomatic restricted collocations.

Although it is not common practice in phraseology to consistently separate “proverb research” from “idiom research”, this distinction must be observed for the purpose of the present study because of the different states of research of the two disciplines. Yet let us first take a short retrospective view on approaches to the problems discussed in this paper.

Matti Kuusi’s important book “Regen bei Sonnenschein. Zur Weltgeschichte einer Redensart” (“Rain in the sunshine. Toward the world history of a saying”) appeared in 1957. From today’s perspective, it is hard to imagine how it was possible at that time, without our modern media, to gather such extensive research material from languages across the world. Kuusi succeeded in recording more than 3,000 variants of the saying mentioned in the title of his book, taken from hundreds of languages and dialects spoken in various countries and continents, and he interpreted them against their particular mythological, cultural and folkloric background. In his introduction, Kuusi (1957: 5–26) describes his surveys, which consisted, above all, of a number of circulars sent to all researchers and research institutes within his reach and an untiring correspondence with scientists and non-professionals all over the globe. Kuusi also tells us about the problems he experienced trying to get into contact with representatives of “difficultly accessible” countries. He was afraid, for example, that several postal items had got lost. Finally, the different alphabets and transcriptions of all the languages in-

volved posed a considerable problem, for which Kuusi again received international support.

As for Kuusi's work (1957), it is a matter of course for many paremiological studies to research the linguistic and historico-cultural parallels of proverbs within their global dimension, cf. e.g. Aquilina (1972), Kuusi (1985) or, more recently, Paczolay (1994, 1997). All these authors present extensive collections of proverbs that are common in many European languages and far beyond. Mieder's bibliography (2004: 259ff) cites a large number of similar proverb studies. These studies are not primarily set in linguistics but in a wide-ranging ethnological-folkloristic and cultural framework. This may be one of the reasons why phraseology research – or to be precise: idiom research in the linguistic sense – never really noticed them. Until recently, idiom researchers have had quite different interests.

Quite probably, a systematic multilingual investigation of idioms, along similar lines as outlined above for proverbs, would produce promising results. Although there are many idioms that would be worth looking at in this respect: with regard to their variants, equivalents and spread across a great variety of languages, such attempts are restricted to a few small studies. There is no tradition of worldwide idiom studies that would be comparable to the prosperous tradition of international cooperation in the field of proverb research, a gap that was not even recognised until recently.

Now and then, idiom researchers have noticed common features of the idioms of some European languages. Menac (1987), for example, presents an inventory of several dozens of – supposedly – common European idioms, drawn from six languages of diverse genetic relationships, namely two Slavonic (Croatian, Russian), two Germanic (German, English) and two Romance (French and Italian) languages. Her analysis reveals quite a large number of near-equivalent idioms in these languages. These cross-linguistic (also called “international”) similarities are attributed exclusively to phenomena of loan translations.

Another comparative study of idioms (Korhonen 1991) covers nine European languages: German, Finnish, French, Italian, English, Swedish, Russian, Hungarian and Estonian. The article is structured synchronically, and it focuses on cross-linguistic equivalences and differences of German and Finnish idioms (cf. the traditionally postulated cross-

linguistic equivalent types labelled as full or total, partial and non-equivalents), whereas the idioms of the other seven languages serve as a control group. More carefully than Menac (1987), Korhonen refers to the origins of the idioms and to the so-called “European cultural heritage”, so that in some cases their common sources and the ways of borrowing from one language to the next become visible. (Admittedly, his results could well be complemented by diachronical studies that take account of the influence of Middle Low German on Swedish and thus also on Finnish idioms, cf. Naumann 1989; Braunmüller 1997, 2004).

Other attempts to analyse idioms across several languages fall far behind the above-mentioned studies. Some research papers rashly label idioms found in a few languages as **internationalisms** or even **universals**. Braun and Krallmann’s article (1990) may serve as one example of several other studies of this kind. The authors report their detection of “international proverbial sayings”, which they assert were unknown at the time of publication. They use the German term **Inter-Phraseologismen** for this phenomenon and illustrate their “discovery” by means of long lists of quite similar idioms in German, English, French, and, in part, Italian.

Mokienko (1998) offers similar lists, supplemented with material from Slavonic languages. Again, idioms of very different origin are presented in a highly unsystematic way. What Mokienko does not mention – and what his lists have in common with those of Braun and Krallmann – is the fact that a number of these idioms depend on well-known textual sources (the Bible, fables, belles-letters, or quotations from prominent people, etc.). Mokienko (1998) calls the tendency of these idioms to converge either internationalism or Europeanism, and in places even universality.

Several other “cross-cultural” studies start from the traditional grouping of idioms into “thematic groups” (e.g. idioms with body part, animal or garment constituents), mistakenly referred to as “onomasiological” approach as well. They come to the result that there are “similar idioms” in some languages. Porsch’s (1997) discovery that “crocodile” – a wild animal not native to Europe – can be found nevertheless in the phraseology of German, Polish and Portuguese, may be regarded as an absurd attempt to cope with this topic (for the spread of the idiom *to shed*

crocodile tears, see examples (4-6) below). Along the same lines, Mejri (2004: 67) classifies the French idiom *jeter le gant* (English *to throw down the gauntlet*, cf. (1) below) as belonging to the thematic group of “garment” (together with coat, shirt, shoes and hat) instead of taking its culture-based nature into account and fully understanding its cultural foundation as its essential feature (the expression refers to a gesture in medieval knighthood, a sign of issuing a challenge).

There have been different attempts at explaining cross-linguistic similarities of idioms. Earlier views to the effect that genetic affiliation could play an important role, as e.g. postulated by O. Nagy (1979) for some similarities between Hungarian and Finnish phraseological units, gave way to the view that the increasing influence of English on other languages is central to the explanation (e.g. Mokienko 1998). However, all the articles mentioned above refer to the somewhat nebulous concept of a “cultural European heritage”, which, for its part, would need a precise explanation. The opinion that similarities between phraseological units can be traced back to the times of the migration of peoples (Piñel 2004: 377) seems to be a special case.

In summary, cross-linguistic similarities of idioms, as opposed to proverbs, have never been studied systematically. Lists of so-called Europeanisms or internationalisms remain accidental; they are merely unmethodical, noncommittal compilations of heterogeneous material. Their main weaknesses are the small number of languages analysed and the complete lack of a theoretical foundation. It is unacceptable to use such terms without the systematic empirical collection of data of languages of different continents and distant cultures, dialects, minor languages, and historical varieties. Without a systematic approach, the discovery of idioms that are spread across various European languages is mere chance, depending on the individual researcher’s accidental foreign language informants or on his/her own knowledge of languages. In view of Kuusi’s (1957) global approach as early as in the fifties, however, it should be even easier today, with the help of our modern digital media, to gather data for at least a large number of European languages.

2. *European idioms: a cross-linguistic and cross-cultural research project*

As terms in idiom research, Europeanism and internationalism lack a clear definition. From a theoretical viewpoint, all the crucial questions as to how relevant Europeanisms/internationalisms can be singled out and determined remain unanswered. Some seem to have never been asked: by what criteria can idioms be defined as Europeanisms/internationalisms? How many languages are to be involved? How great do the similarities have to be, and what metalanguage should be used? Further central issues include the questions of how many Europeanisms/internationalisms exist in every single language, or in all of them collectively, and to which cultural domains they belong.

A short time ago, I have set about investigating more systematically into the question of idioms that are supposed to be widespread across many – mainly European – languages. The principal goal of my envisaged research project will be the creation of an inventory of factually existing “widespread European idiom types”. This task requires substantial empirical preliminary work, which cannot be carried out by one single researcher. An international, multi-language cooperation of linguists will be necessary to analyse the similarities of idioms in as many languages as possible. A prejudiced “Eurocentric” view must be avoided: although it is true that the idioms focused upon are mainly taken from European languages, this does not mean that they do not occur in non-European languages (see section 4 below). Thus, for idioms supposedly widespread across European languages and beyond, we suggest the term **Widespread Idiom** (or **WI** for short) as a working term.

To single out possibly widespread WIs, the first step was to subdivide the total number of idioms into smaller and therefore comparable entities (cf. Piirainen 2004a, in print a, b). Idioms were categorised according to their cultural foundations, using the typology of principal cultural phenomena that underlie idioms (cf. Dobrovol’skij/Piirainen 2005: 205ff). This typology was guided by the multi-language idiomatic material itself. A large number of idioms had been analysed to learn what types of cultural knowledge play a role in motivating the links between literal and figurative meanings. Starting exclusively from this linguistic basis, one can note that there emerge structures of cultural phenomena that are strikingly similar to those established

by semioticians attempting to define the notion of “culture”, namely categories such as social, material and mental culture. Since language as a whole is a mental phenomenon, the category “mental culture” is unsuitable for describing linguistic phenomena and had to be replaced by further subcategories (i.e. textual dependence of idioms, fictive conceptual domains and symbolic knowledge as the principal subcategories of cultural foundation). Let us illustrate this with some examples.

One type of cultural foundation is represented by idioms whose underlying cultural knowledge chiefly goes back to knowledge of **aspects of social culture**. This type can be divided into smaller subgroups. One of them consists of semiotised gestures as source concepts of idioms. It does not matter whether the gesture in question is actually performed or not. The widespread idiom *to tear one’s hair out* will be discussed in section 4, example (11). The gesture of throwing the gauntlet mentioned above belongs to this group of cultural foundation as well (rather than simply to a ‘garment’ group). The first criterion for being supposedly “widespread” is if an idiom occurs not only in some Central European languages such as English, German or French, but also in some geographically more distant and genetically unrelated languages. Idiom (1) would be such a candidate. (See below for abbreviations of languages.)

- (1) Eng *to throw down the gauntlet*
 Grm *jmdm. den Fehdehandschuh hinwerfen*
 Fr *jeter le gant à qn*
 Spn *arrojar el guante a al*
 Rus *бросать/бросить перчатку кому-либо* “to throw the glove to sb.”
 Fin *heittää jklle taisteluhansikas* “to throw the battle-glove to sb.”
 all meaning ‘to issue a challenge, to challenge sb. to a contest’

Sometimes other aspects of social culture overlap with aspects of material culture as source concept. Clearly determinable subgroups include, for example, the domains ‘sport’, ‘games’ or ‘theatre and music’. The following idioms taken from these subgroups can be found in several languages and are potential WI candidates: *below the belt*, *to play one’s trump card*, *behind the scenes*, *to play a role* or *to play the first/second fiddle*.

Another type of idioms is based on source concepts that can be fully ascribed to **aspects of material culture**. These aspects can be divided into a number of smaller subgroups as well, extending from knowledge of everyday life, e.g. modern traffic, telecommunication or urban dwelling (see examples (7) and (10) in section 4), and other cultural achievements (such as writing with black ink or print products) to artefacts of craft or agriculture in earlier days. Examples (2) and (3), drawn from such subgroups of material culture, may be WI candidates.

- (2) Eng (*to have something down*) *in black and white*
 Fr (*écrire quelque chose*) *noir sur blanc*
 Grm (*etwas*) *schwarz auf weiß* (*haben*)
 It (*mettere*) *nero su bianco*
 Fin (*jotain on*) *mustaa valkoisella* “(sth. is) black on white”
 all meaning ‘(something is) official, in writing or print’
- (3) Eng (BrE) (*to be*) *the fifth wheel (on the coach)*; (AmE) *to be the third wheel*
 Grm *das fünfte Rad am Wagen* (*sein*)
 Fr (*être*) *la cinquième roue du carrosse*
 Fin *olla kolmantena pyöränä* (*vaunuissa*) “to be (like) the third wheel (on a/the wagon)”
 all meaning ‘(to be) someone who is with a group of people, even though that group may not want or feel comfortable with them’

The (almost certainly) largest type of idioms suspected to constitute WIs can be subsumed under the label **textual dependency**. Originally, these are quotations or allusions, and so they refer to an identifiable textual source. Subdividing idioms of this type into smaller groups is easy because their textual origins offer a suitable structure. This type includes a large number of idioms that can be traced back to narratives, tales, works of classical literature or quotations from prominent people (*with seven-league boots*; *to tilt at windmills*; *cannot see the wood for the trees*; *to build castles in the air* etc.), to fables and classical antiquity (*the lion’s share*; *to pull the chestnuts out of the fire for sb.*; *to bell the cat*; *a snake in the grass*) or to the Bible, which make up the largest group of all supposedly widespread idioms. All of the following idioms are proper candidates for widespread idioms: *to cast pearls before swine*; *to swim against the tide/stream*; *to take someone under one’s wing*; *to separate/sort the wheat from the chaff*; *(to be) built on sand*; *to pour out one’s heart to someone*; *not to hide one’s light under*

a bushel; the scales fall from someone's eyes; someone's hair stands on end; by the sweat of one's brow or to wash one's hands of something. As substantial studies of biblical idioms have already been carried out for various languages, the search for WIs could start directly from this subgroup.

Among the remaining types of cultural foundation, there is no such abundance of WI candidates: neither idioms that can be traced back to **fictive conceptual domains**, such as superstition, ancient folk theories or other pre-scientific conceptions of the world (e.g. *to thank* *one's lucky stars; hell on earth*), nor idioms dealing with **cultural symbols** (e.g. *to look on the black side, to have a heart of gold*) can be supposed to provide many WIs.

On the other hand, some other domains that do not belong to the typology of the cultural foundations of idioms need to be considered here. The complexes of 'nature and animals' behaviour' (e.g. *to break the ice; to show one's teeth to a person; to play cat and mouse with sb.*) and 'somatic phenomena' (e.g. *from head to foot; to be all ears; to open someone's eyes; (nothing but) skin and bones; to the last breath*) seem to contain widespread idioms.

Cross-linguistic comparisons of these smaller subgroups revealed about 150 idioms that can supposedly be found in various European languages. The next step consisted in designing questionnaires (according to the target group, drawn up in English, German, French or Russian) and to send them out as e-mail files to as many idiom research experts as possible. This step met with wide response.

Up to the present day (February 2005), I have received idiom data of the following European languages: Albanian, Bulgarian, Catalan, Croatian, Crimean Tatar, Danish, Dutch, English, Estonian, Faroese, Finnish, French, North Frisian, Galician, Georgian, German, Greek, Hungarian, Icelandic, Italian, Latvian, Lëtzebuergesch, Ladin, Lithuanian, Mordvin Erza, Norwegian, Polish, Portuguese, Romanian, Romansh, Russian, Slovakian, Slovene, Spanish, Swedish, Swiss German, Turkish, Ukrainian and Upper-Sorbian.

Functioning as a contrast group are the non-European languages Chinese, Egyptian Arabic, Japanese, Korean and Thai. Nevertheless, data from more than hundred European languages are still lacking, as the following overview will show.

3. The languages of Europe

3.1. Language families and idiom research

For the purposes of the research project envisaged here, i.e. the investigation of idioms that are widespread across European languages, it seems helpful to gain an overall view of the languages spoken in this continent and take stock of the idiom research that has been carried out so far for these languages. Attempts to demarcate “Europe” from other cultural and linguistic areas by means of distinctive features are doomed to failure. Culturally, it is impossible to establish clear boundaries. Rather, the so-called European Occidental culture has its roots in the Middle East as well as in North Africa. The linguistic situation is quite similar. One of the major language families spoken in Europe, often called *Indo-European*, has substantially more representatives in Asia than it has in Europe. In the Indian subcontinent alone, we find much more Indo-European (i.e. Indo-Iranian) languages – and much more speakers of these languages – than in the whole of Europe. Moreover, a number of languages do not “stop” at any virtual border within the Eurasian continuum, e.g. Turkish, Samoyed and various Altaic and Caucasian languages.

Therefore, **Europe** will be understood here simply in terms of its geographic boundaries, extending eastwards as far as the Ural mountain range and – as is usual in linguistic typology – including the Caucasus region. According to this strictly geographical definition of Europe, more than 140 languages should be taken into account; this number does not include dialects and diaspora languages (Yiddish, Judeo-Spanish or Romani). For simplicity’s sake, let us ignore, as far as possible, the problem as to where to draw the line between “languages” and “dialectal varieties”; cf. Crystal’s (2000: 11ff) discussion of how to define a language and how to distinguish between language and dialects. Following below is a list of European languages, classified according to their genetic affiliation. It turns out that the linguistic map of Europe is made up of at least one totally isolated language, namely **Basque**, and six completely unrelated phyla (**Indo-European**, **Uralic**, **Altaic**, **Eskimo-Aleutian**, **Semitic** and the **Caucasian languages**).

At the same time, let us take a glance at the state of idiom research within these languages. Until now, much of idiom research has been concerned with a few standard languages, at first with Russian, Ger-

man, French, Finnish, Hungarian as well as Czech, Slovak and Polish, but fairly recently also with Romanian, Bulgarian, Slovene, Spanish, Dutch, Icelandic, and Danish, and in its infancy also with Greek, Italian, Lithuanian and Latvian. Only one minor language, namely Lëtzebuergesch, has been studied thoroughly with respect to its idioms (Filatkina 2005). Dialects, apart from one Low German dialect, have been ignored completely. Cf. Burger et al. (in print); Piirainen (2000, in print c) for more details.

In summary, only two of the European phyla (namely Indo-European and Finno-Ugrian, a subfamily of the Uralic languages) are sufficiently represented in the literature. The majority of European languages have hardly ever or even never been the object of idiom research. Apart from at best two dozens of easily accessible languages in which more or less extensive idiom research has been carried out, no idiom data are available at all, neglecting a large number of European languages (in fact more than 80% of the languages spoken in Europe). Things are largely different with regard to proverb studies since paremiology has been practised on a more comprehensive, multi-language scale from the very beginning. There are a great number of dictionaries available for dialectal proverbs of Europe. The situation of minor languages, too, is much better in this area, cf. e.g. the comprehensive proverb dictionaries of Maltese (Aquilina 1972) or Sorbian (Hose 1996).

3.2. Overview of the languages spoken in Europe¹

The Indo-European language family

The number of living Indo-European languages spoken in Europe is estimated at about 60. Traditionally, these languages are divided into three almost isolated languages (*Albanian*, *Armenian* and *Greek*) and five or six larger branches, whose members show closer linguistic affinities with each other: these are the *Germanic*, *Romance*, *Celtic*, *Baltic*, *Slavonic* (or the latter two seen as one branch, *Balto-Slavic*) and *Indo-Iranian* languages. To date, no idiom data have been available for

¹ The following details are mainly taken from Price (1998), van der Auwera (1998: 6-14) and Haarmann (1993, 2002); for Basque, Faroese, Galego, Karaim, Maltesian, Raeto-Romance and Welsh see also the respective articles in Stolz (2001).

three of these Indo-European branches, namely Armenian, Celtic and Indo-Iranian. A traditional classification within these branches would look as follows:

- The *Germanic languages* are composed of the *North Germanic group* (*Icelandic, Faroese, Norwegian, Swedish, Danish*) and the *West Germanic group*: *Dutch, German, English, Frisian* (with its *West-, North- and East-Frisian varieties*) and *Lëtzebuergesch* or *Luxembourgeois*, the latter being a minor language and one of the official languages of Luxembourg, whereas *Swiss German*, in spite of the comparable diglossic and socio-linguistic circumstances, is not labelled as a language but as a dialect).
- The *Romance languages* are divided into *Sardinian*, spoken on the island of Sardinia (while *Corsican* is seen as an Italian dialect) and the *Continental Romance languages*. The latter consist of one western group, with *French, Occitan* (or *Provençal*), *Italian, Friulian, Ladin*, and *Romansh* (whose main dialects are known as *Sursilvan* and *Sutsilvan*, which together make up Romansh, one of the official languages of Switzerland, also subsumed under *Raeto-Romance*) as well as *Spanish, Catalan, Galego* (or *Galician*, an official language in the autonomous region of Galicia) and *Portuguese*, and one eastern group, consisting of *Romanian* and *Aromanian*, also known as *Macedo-Romanian*, which is sometimes seen as a distinct language and sometimes as a dialect of Romanian.
- All surviving members of the *Celtic languages* belong to the so-called *Insular Celtic group*, also known as *Brittonic* or *Brythonic languages*, with the subgroups of *Irish, Scottish Gaelic* (or *Scots Gaelic*, spoken along the northwest coast of Scotland and in the Hebrides islands) and *Welsh*, which is in a relatively secure position in Wales. *Breton*, too, spoken by a small minority in Brittany in north-western France, counts among the *Insular Celtic languages*, because of the close relationship it has with them.
- There are only two *Baltic languages*, *Latvian* and *Lithuanian*.
- The *Slavonic languages* are subdivided into three groups: the eastern group consists of *Russian, Belorussian* and *Ukrainian*, the western

group of *Czech, Slovak, Polish* along with the minor languages *Sorbian, Kashubian* and *Polabian*, whereas the southern group comprises the languages of *Slovene, Croatian, Serbian, Bosnian, Bulgarian* and *Macedonian*.

- Most of the *(Indo-)Iranian languages* are located in Asia. Only the languages *Kirmanji, Talysh, Tati*, and *Ossetic* (located in the Caucasus region) belong to the European area.

The Uralic language family²

According to van der Auwera (1998: 8), seventeen Uralic languages are spoken in Europe. They fall into two branches, the ***Samoyed group*** (of which only one language, *Tundra Nenets*, can be found in Northeast Europe) and the main group of ***Finno-Ugrian languages***.

The ***Ugric group*** of the Finno-Ugrian phylum is composed of *Hungarian*, whose idioms have been studied very well, and the *Ob-Ugric languages*. Of the latter group, only *Mansi* is located in Northeast Europe. There are no known idiom collections for these languages.

As far as the *Finnic group* is concerned, there is no full consensus as to whether they should be divided into three or four sections. A common subdivision might look as follows:

- The *Permic group* comprises the languages *Udmurt*, also known as *Votyak*, and *Komi* or *Zyrian/Zirene* (with the major groups of the *Komi-Zyryan*, the *Permyaks* and the *Komi-Yazua*).
- The *Volgaic group* (or *Volga Finnic group*) consists of the languages *Mari* (formerly *Cheremis*) and *Mordvin*, which is one of the largest Uralic languages in number of speakers, ranking after Hungarian and Finnish and/or Estonian. Mordvin has two major dialects, which are sometimes considered to be separate languages: *Moksha*, spoken in the west of Mordvinia, and *Erza* or *Erzya* (which is represented in this study), spoken in the eastern parts.

² Details are taken from Abondolo (1998), for the Finnic languages in particular also from Dahl/Koptjevskaja-Tamm (2001).

- The *North Finnic* or *Baltic-Finnic group*. This group includes the languages *Ingrian*, *Votic* and *Livonian*, which are considered to become extinct within the next generations. This group also includes the less used languages *Vepsian* and *Karelian* (with the varieties *Olonets* and *Ludic*, considered to be a blend of Karelian and Vepsian), as well as the languages *Finnish* and *Estonian*.
- The origin of *Sámi*, indigenous to the area of Lapland and other northernmost parts of Europe, is uncertain. Whether the Sámi varieties form a separate group or must be subsumed under the North Finnic group, remains a matter of controversial discussion.

Of the Finnic group of languages, Finnish idioms have been studied most thoroughly. Rich data are also available for Estonian idioms, among other things through the dictionaries by Õim (1993) and Hanko/Liiv (1998). Furthermore, there is quite a substantial documentation of Karelian idioms (Федотова 2000). As far as the remaining eleven Finnic languages – belonging to the European area as well – are concerned, it would be quite difficult to obtain reliable data on their idioms. Again, for some of these languages, the study of proverbs (Komi, Livonian, Vepsian and Votyak) has a long tradition, cf. Kuusi (1985).

Basque

Basque is the only European language that is completely isolated. It is regarded as the only remaining language of those that were spoken in southwestern Europe before they were driven back by the Romance languages. Basque with its standard variety and six dialects is still in a stable position in the Euskari province in Spain. A collaboration of native speakers of Basque would be desirable for the project described here.

The Semitic language family

Maltese, the national language of the Maltese Islands, belongs to the southern central group of the Semitic languages. Although Maltese is regarded as a mixed language, Arabic is its genetic basis. It is closely related to the Arabic dialects of Tunisia and Algeria. Maltese is the only form of Arabic that is written in the Latin alphabet. Except for proverbs (Aquilina 1972), no phraseological material has been made available. A collaboration of an expert of Maltese has been announced.

There are two further, minor Semitic varieties spoken in Europe: **Cypriot Arabic**, spoken by the inhabitants of the village of Kormakiti in Cyprus, and **Assyrian**, some speakers of which live on the European side of CIS states (Price 1998: 110, 20).

The Eskimo-Aleut language family

Greenlandic, one of the official languages of Greenland and spoken by a small group of Greenlanders in Denmark, belongs to the *Inuit* or *Eastern Eskimoan* branch of the Eskimo-Aleut language family. A collaboration of an expert of Greenlandic on the envisaged project has been announced.

The Altaic language family

Haarmann (2002: 43) identifies 65 Altaic languages. About twelve of these are located to the west of the Ural mountain range (the virtual eastern border of Europe). These languages can be divided into the subfamilies of the *Mongolian* and the *Turkic languages*. The only representative of the Mongolian branch in Europe is the language of *Kalmyk* or *Kalmuk*, spoken in a region to the west of the lower Volga, northwest of the Caspian Sea. This language belongs to the *Oirat* (or *Oyrat*), i.e. the western branch of the Mongolian language group.

Among the Turkic languages on the European side, we can first subdivide the so-called **Common Turkic languages**. They consist of the north-western or *Kipchak* branch (with *Bashkir*, *Karachi-Balkar*, *Kumyk*, *Crimean Tatar*, *Tatar* and *Karaim*, the latter belonging to the north-western or *Kipchak* branch), a southern group (with *Turkish*, *Gagauz* and *Azerbaijani*, the latter belonging to a south-western subgroup, also called *Oguz* or *Turkmen*), and a central group (*Nogai* or *Noghay*). Secondly, there is **Chuvash**, a separate branch of the Turkic language; it differs considerably from the Common Turkic languages.

Thus, besides *Turkish*, spoken by a number of people in the Balkans, e.g. in Bulgaria, and in Turkey, several other Turkic languages can be found in Eastern Europe. Geographically, they are spread from Lithuania to the south-easternmost areas of Europe: a last group of *Karaim* speakers live near Vilnius in Lithuania and in south-western Ukraine, where it is also moribund. *Bashkir* is spoken in Bashkortostan, and *Karachi-Balkar* and *Kumyk* in the north of the Caucasus Mountains. *Tatar* is spoken in the republic of Tatarstan in west-central Russia and

in Romania and Bulgaria. *Kazan Tatar* is the major Tatar dialect and a literary language. *Crimean Tatar* is most closely related to *Kumyk*, not to *Kazan Tatar*. There are still speakers of this language, living as a minority in Ukraine and Romania. *Gagauz* is located in regions to the north-west of the Black Sea, in Ukraine, Moldova and the Balkans, whereas *Azerbaidjani* and *Nogai* are spoken in the northern and the southern Caucasus areas, respectively. *Chuvash* is spoken mainly in Chuvashia and nearby regions south of the middle course of the Volga.

Although phraseological research has been undertaken on some of the Turkic languages mentioned here (cf. the bibliography of Бушуй 1987, and several proverb dictionaries, e.g. Yurbasi 1993, 1996), these studies are hardly accessible. For most of these languages, it will be very difficult to gain data on idioms. We succeeded in winning speakers of Turkish and Crimean Tatar over as informants for the envisaged project.

The Caucasian languages³

More than three dozen languages spoken in the quite small territory north and south of the Greater Caucasus Range and its foothills are to be classified as “the Caucasian languages” (i.e. the autochthonous languages, not to be confused with “languages spoken in the Caucasus region”, cf. the above-mentioned Armenian, Kalmyk, and several Iranian and Turkic languages; see Comrie 2005: 1). The usual division of the Caucasian languages into three typologically different phyla is named after their geographic locations. These families are the *North East Caucasian*, the small *North West Caucasian* and the *South Caucasian languages*. The exact genetic affiliations of all Caucasian languages are still unclear to some extent, with respect to both the relationships between the three major groups and some internal categories. In the following, we restrict ourselves to a rough outline.

- The *North East Caucasian languages* (also known as *Nakh-Daghestanian*) fall into the *Nakh group* (with the fairly similar languages *Chechen* and *Ingush* and a very small language called *Bats*

³ The following details are taken from Boeder (2005), Hewitt (2005), and van den Berg (2005).

or *Tsova-Tush*) and the *Daghestan subfamily*. The latter is divided into the *Avar-Andic-Tsezic* group (including the languages *Andi*, *Akhvakh*, *Avar*, *Bagvalal*, *Bezhta*, *Botlikh*, *Chamalal*, *Godoberi*, *Hinukh*, *Hunzib*, *Karata*, *Khvarshi*, *Tindi* and *Tzez* or *Dido*), the *Lezgian group* (including the languages *Agul*, *Archi*, *Budukh*, *Khinalug*, *Kryts*, *Lezgian*, *Rutul*, *Tabasaran*, *Tsakhur* and *Udi*), and the *Lak-Dargwa group*. Several languages of this family are close to extinction.

- The *North West Caucasian languages*, also known as *Abkhazo-Adyghian*, consist of three branches: *Abkhaz-Abaza* (with the two very similar languages *Abaza* and *Abkhaz*), the *Circassian group* (including *Adyghe* and *Kabardian*) and the extinct *Ubykh*.
- The *South Caucasian languages* (also called *Kartvelian languages*) are comprised of *Georgian*, *Svan* and the *Zan group* (consisting of *Laz* and *Mingrelian* or *Megrelian*).

As is the case with the Turkic languages, it will be very difficult to gather reliable data on the idioms of most of the Caucasian languages. Although “[t]he Kartvelian languages are one of the best studied language families of the world” (Boeder 2005: 77), there is not much information on their idioms available. Among the Caucasian languages, Georgian has the longest written tradition; much linguistic research on Georgian has been carried out. Linguists also dealt with Georgian proverbs, and large collections of them exist (cf. Boeder 1985, 1991). To our knowledge, however, studies on Georgian idioms are only very small. The trilingual phraseological dictionary by Gamrekeli et al. (1973) seems to be largely obsolete.

4. Examples of questionnaire data and first results

This section will discuss four examples that illustrate different problems of widespread idioms. Each example starts with a full presentation of the idiom data as given by the native speakers questioned. These data will be arranged according to the language families in question, in line with the overview given in the previous section. This arrangement has been chosen for the sake of clarity; it does not imply that the genetic affiliation influences the spreading of the idioms. Sorting them accord-

ing to geographic location (e.g. from the languages spoken in North West Europe to those spoken in the southeast) would have been less clear. If a given idiom type does not exist in a particular language, it is marked by (-).

After this basic data, three topics will be treated: (i) the **distribution** of the WI across the languages (as far as currently available with the aid of our informants), (ii) potential special features of the **inner form** of individual idioms (of their morphosyntactic and lexical structure, i.e. at the level of the “literal reading”) and (iii) potential peculiarities regarding the **phraseological meanings** (also called “figurative” or “actual meanings”). Theoretical reflections on the nature of WIs, however, must be reserved until further empirical data have been analysed.

4.1. “European idioms” or “widespread idioms”?

The idiom *to weep/shed crocodile tears* ‘to show sadness that is not sincere, to display insincere grief’ (4) is a suitable example of our data. This idiom is a genuine case of textual dependence. It originates from old legends, according to which crocodiles weep like a child and shed tears in order to lure their victims and when they devour them. In his influential “Buch der Natur” (“Book of Nature”, 1350), Konrad von Meigenburg describes the crocodiles’ behaviour “wenn ez aines menschen ertoett, sô waint ez in” (when it kills a person, it weeps over him) (Bächthold-Stäubli 1933: 596). This knowledge about crocodiles – and accordingly, of the conventionalised idiom type – is believed to be of classical origin and to have been spread across the European languages by the humanists, mainly through Erasmus’ influential “Adages” and Luther’s works. The data compilation looks as follows:

(4) TO WEEP/SHED CROCODILE TEARS

Indo-European languages**Germanic**

Ice	<i>gráta/skæla krókódilatárur</i>
Far	(see (6a))
Nor	<i>gråte krokodilletårer</i>
Swd	<i>gråta krokodiltårar</i>
Dan	<i>græde krokodilletårer</i>
Dut	<i>krokodillentranen huilen</i>
Grm	<i>Krokodilstränen weinen/vergießen</i>
Eng	<i>to weep/shed crocodile tears</i>
NFrS	<i>Krokodilstuaren skruål</i>
LëtZ	<i>Krokodilstréine kräischen</i>
SwG	<i>Krokodilsträne brüele</i>

Romance

Fr	<i>verser des larmes de crocodile</i>
It	<i>versare lacrime di coccodrillo</i>
Rom	<i>a plînge cu lacrimi de crocodil</i>
Spn	<i>derramar/llorar lágrimas de cocodrilo</i>
Ctl	<i>plorar amb llàgrimes de cocodril</i>
Glc	(-)
Prt	<i>chorar lágrimas de crocodilo</i>
Rmns	<i>criðar larmas da crocodil</i>

Baltic

Ltv	<i>raudāt krokodila asaras</i>
Lith	<i>krokodilo ašaras lieti/krokodilo ašaromis verkti</i>

Slavonic

Rus	<i>(лить/проливать) крокодиловы слезы</i>
-----	---

Ukr	<i>лити/проливати крокоділячі слези</i>
Slva	<i>ronit' krokodilie slzy</i>
Pol	<i>ronić krokodyle łzy</i>
Usrb	(see (6b))
Slve	<i>točiti/pretakati krokodilje solze</i>
Cr	<i>roniti krokodilske suze</i>
Blg	<i>роня/проливам крокодилски сълз от плача с крокодилски сълзи</i>

Albanian *derdh lot krokodili***Greek** *χρῖνω κροκοδείλια δάκρυα***Finno-Ugrian languages****Ugric**

Hng	<i>krokodilkönnyeket hullat/ejt</i>
-----	-------------------------------------

Finnic

MrdE	<i>pravtoms krokod'il'en's'el'ved'ŷ</i>
Fin	<i>itkeä/vuodattaa krokotiilin kyyneleitä</i>
Est	<i>krokodillipisaraid valama</i>

Turkic languages

Trk	<i>timsah gözyaşları dökmek</i>
CrT	(-)

Caucasian languages

Grg	<i>niang-is creml-eb-i</i> “crocodile’s tears”
-----	--

(i) Idiom type (4) can be found in almost all the languages analysed here. Furthermore, the idiom occurs in various languages outside of Europe. The following examples (5) provide an identical word string “crocodile tears”, equally meaning ‘insincere, hypocritical tears’ (different word strings for the same denotatum, e.g. Japanese *sora namida* “empty tears” or Korean *nun-mul-eul ssa-da* “wring out the tears” are not significant in this context):

(5a) Egyptian Arabic *dumū' timsah*

(5b) Thai *namta jorrake*

(5c) Chinese *liú xià/liú chū*

So far, there has been no research on the way this idiom type has taken to get into the Asian languages – whether it is to be traced back to

common narrative traditions dating from antiquity (as has been shown to be the case with many idioms originating from fables and subsequently passed on via the “Panchatantra”) or whether it is based on a more recent layer of borrowing.

(ii) The idioms in (4) are almost identical lexically, literally meaning “weep/shed crocodile tears/tears of a crocodile” or (Ctl and Rom) “weep with tears of a crocodile”. Only the verb (as given by our informants) varies somewhat between ‘to weep/cry’ and ‘to shed’ – some informants gave both verbs. All other differences, at the morphosyntactic level, are language-internal, due to the typological character of the particular language in question. The more synthetic languages will use a compound, such as the Germanic languages (Grm *Krokodilstränen*), or Hng *krokodilkönnyeket* and Est *krokodillipisaraid*. Other languages, especially the Romance ones, make use of genitive constructions (e.g. Spn *lágrimas de cocodrilo*). Most of the Slavonic languages provide adjective-noun constructions, cf. Rus *крокодиловы слезы*. Obviously, such differences do not concern the aim of singling out WIs. Thus, the languages analysed here show a unified picture. Only two of the minor languages go their own way: Faroese uses the word for ‘horse’ instead of ‘crocodile’ (6a), while Upper-Sorbian adapts the ‘wolf’, well-known for its “bad” behaviour or as a symbol of malice and evil. The different passive construction should be noted here (6b).

(6a) Far *gráta rossatár* “weep horse-tears”

(6b) Usrb *wjelče sylzy ronja někomu* “wolf’s tears are running down sb.’s face”

(iii) The phraseological meaning of all the idioms presented here is stable and does not require further comment.

These examples have several theoretical consequences. They demonstrate that the proper place of my research project is not in contrastive linguistics (in the traditional sense). Rather, the project is aimed at either corroborating or refuting the oft-claimed uniformity of European idioms, whereas the relationship of equivalence between idioms of different languages is irrelevant. Similarly, it is not concerned with the question of whether individual idioms are equivalent with respect to their textual behaviour or all their diasystematic features, e.g. their familiarity. To the contrary, research should even include obsolete idioms as well, at least initially. The idioms in (6) will surely need to be discussed

further: we need to ask whether they belong to the same idiom type. Moreover, the occurrence of this idiom type in several Asian languages reinforces my call for the inclusion of non-European languages into the project. It may be more appropriate to speak of a “widespread idiom type” here than of a “European idiom type” proper.

4.2. A gap in the geographical distribution

Let us now look at an example based on an aspect of modern material culture. ‘Radio communication’ is the source concept underlying the following idiom type (7). The idioms refer to two or more people having similar ideas and opinions, understanding each other very well, *as if* they were on the same frequency, like broadcasting transmitter and receiver. The following data are currently available:

(7) TO BE ON THE SAME WAVELENGTH AS SOMEONE

Indo-European languages		Baltic	
Germanic		Ltv	<i>būt uz viena viđņa</i>
Ice	<i>að vera á sömu bylgjulengd (og einhver)</i>	Lith	(-)
		Slavonic	
Far	<i>vera á bylgjulongd (við ngn)</i>	Rus	(-)
Nor	<i>være på bølgelengde (med noen)</i>	Ukr	<i>бути на одній хвилі</i>
Swd	<i>vara på samma våglängd (med ngn)</i>	Slva	(-)
Da	<i>være på bølgelængde (med ngn)</i>	Pol	(-)
Dut	<i>op dezelfde golflengte zitten (met iemand)</i>	Usrb	(-)
Grm	<i>(mit jmdm.) auf der gleichen Wellenlänge liegen</i>	Slve	<i>biti na isti valovni dolžini</i>
Eng	<i>to be on the same wavelength (as sb.)</i>	Cr	<i>biti na istoj valnoj duljini</i>
NFRs	<i>üp di salev Welenlengdi wiis</i>	Blg	(see (9))
Lëtz	(-)	Albanian (-)	
SwG	<i>uf dr gliiche Wällelengi sii</i>	Greek	<i>εκπέμπουν στο ίδιο μήκος κύματος</i>
Romance		Finno-Ugrian languages	
Fr	<i>être sur la même longueur d’onde</i>	Ugric	
It	<i>essere sulla stessa lunghezza d’onda</i>	Hng	<i>egy/azonos hullámhosszon van vkivel</i>
Rmn	(-)	Finnic	
Spn	<i>estar en la misma onda</i>	MrdE	(-)
Cat	<i>estar en sintonia (amb algú)</i>	Fin	<i>olla samalla aaltopituudella (jkn. kanssa)</i>
Glc	(-)	Est	<i>samal lainepikkusel olema</i>
Prt	<i>estar na mesma sintonia</i>	Turkic languages	
Rom	<i>a fi pe aceeași lungime de undă</i>	Trk	<i>aynı frekansa sahip olmak</i>
		CrT	(-)

(i) The spreading of this idiom type presents a picture that is quite different from that of the previous example (4). If the data is projected onto a geographic map, we can note an uninterrupted gap for the East European languages (Lithuanian and several Slavonic languages), running like a wedge between the northern and the southern areas of distribution. In theoretical terms, one has to consider how many European languages are to be involved to constitute a “European idiom type”. On the other hand, there are two quite similar idioms in non-European languages (8), both of which are modern and colloquial.

(8a) Japanese *hachou-ga au* “have the same wavelength”

(8b) Korean *ju-pa-su-ga matt-da* “sb. has the same wavelength as sb.”

(ii) No particular emphasis should be laid on the morphosyntactic structures: types such as “be/lay on the same wavelength (with sb.)” and “have the same wavelength (as sb.)” are used side by side in some languages (cf. the German variant *die gleiche Wellenlänge haben*). Some lexical variants should be noticed as well. The word used for “same” can be omitted (as in Far, Nor and Dan), just as the element meaning “length” (cf. Spn, Ltv and Ukr). Similar to “wavelength” are Turkish *frekans* or Catalan and Portuguese *sintonia*.

(iii) Whereas all the idioms listed in (7) and (8) have the same phraseological meaning (as described above), the seemingly similar Bulgarian idiom in (9) reveals a different semantic structure.

(9) Bulgarian *на същата вълна съм* “I am on the same wave” ‘I busy myself with the same problems or with similar ideas (as sb.)’

The antonymous idiom is just as frequent: *на друга вълна съм* “I am on the other wave”, meaning ‘I occupy myself with different problems, I am thinking about something very different’. In view of the idioms listed above (7-8), the Bulgarian idiom (9) may be called a “false friend” (see section 4.3).

4.3. Phraseological false friends

The next example is from the domain of material culture as well. Housing belongs to the most elementary requirements of human beings. There is a wide range of cultural diversity when it comes to the way that people build their homes. It is astonishing that the conceptualisations of ‘house’ or ‘home’ as source frames manifest themselves in very similar

(cf. Piirainen 2000: 200-244; Dobrovolskij/Piirainen 2005: 187-203).

(ii) We will not elaborate on structural differences between the idioms (e.g. the adverbial use of the Danish example). Lexical differences, however, should be noticed. Apart from the different prepositions and postpositions (*in, between, inside of*), verbs (*be, sit, live*) and the occurrence of an adjective for ‘one’s own’, there are different concepts of the walls: some languages differentiate between the ‘wall of a room’ and an ‘external wall’ (e.g. Grm *Wand* vs. *Mauer* or Fin *seinä* vs. *muuri*), while other languages do not make this lexical distinction. All these facts can influence the meanings of the idioms (cf. (iii)).

(iii) The phraseological meanings of the idioms listed in (10) are not congruent. Thorough semantic analyses are required for every single language. We can observe two main meanings: ‘(to be) in the privacy of one’s (own) home, at home, in an atmosphere of confidence’ on the one hand (mainly in the Germanic languages, but also in Finnish, Greek and Turkish), and ‘to always stay at home, never go out, be isolated, be bored, feel hemmed in’ on the other (mainly in the Romance languages, but also in Lithuanian). The meaning ‘inside a building (of any kind)’ seems to predominate in Icelandic. In some cases one can speak of genuine “phraseological false friends” (cf. Piirainen 2004b: 158f); in other cases one can find different meanings side by side in the same language. From a theoretical viewpoint, the question arises how great the semantic similarities need to be in order to speak of the same idiom type.

4.4. Common source or independent origin?

The next example is taken from the cultural domain of semiotised gestures. The meaning of the following idioms (11) can be roughly paraphrased as ‘to be in deep despair; to be frustrated or worried’ (cf. (iii) below).

(11) TO TEAR ONE'S HAIR OUT

Indo-European languages**Germanic**

Ice	<i>að reyta/rífa hár sitt</i>
Far	(-)
Nor	<i>å rive seg i håret</i>
Swd	<i>slita sitt hår</i>
Dan	<i>rive sig i håret</i>
Dut	<i>zich de haren uit het hoofd trekken</i>
Grm	<i>sich die Haare raufen</i>
Eng	<i>to tear one's hair out</i>
NfFr	<i>hat es to Hiir üttriwen/em kür jen di Hiir üttriiv</i>
LëtZ	<i>'t ass fir sech d'Hoer auszerappen</i>
SwG	<i>sich d' Hoor raufe</i>

Romance

Fr	<i>s'arracher les cheveux</i>
It	<i>strapparsi i capelli</i>
Rmns	<i>sa strair per ils chavels</i>
Spn	<i>tirarse de los pelos</i>
Ctl	<i>estirar-se dels cabells</i>
Glc	(-)
Prt	<i>arrancar-se os cabelos /descabelar-se</i>
Rom	<i>a-și smulge părul din cap</i>

Baltic

Ltv	<i>plēst matus</i>
Lith	<i>rautis plaukusi</i>

Slavonic

Rus	<i>рвать (на себе) волосы</i>
Slva	<i>trhat' si vlasy</i>
Pol	<i>rwać/drzeć włosy z głowy</i>
Usrb	<i>sebi wlosy torhać</i>
Slve	<i>puliti si lase</i>
Cr	<i>čupati si kosu</i>
Blg	<i>скупя си косите</i>

Albanian *shkul flokët (nga marazi)***Greek** *τραβώ τα μαλλιά μου***Finno-Ugrian languages****UgriC**

Hng	<i>a haját tépi</i>
-----	---------------------

Finnic

MrdE	<i>raz'd'ems čer't'</i>
Fin	<i>raasta/repää hiukset päästään</i>
Est	<i>juukseid katkuma</i>

Turkic languages

Trk	<i>saçını başını yolmak</i>
CrT	<i>башыны-сачыны юлкмакь</i>

Caucasian languages

Grg	<i>tma-s igležs</i>
-----	---------------------

(i) The distribution across the European languages shows a quite complete and unified picture here. The same holds for the non-European idioms in (12), which display almost identical word strings and phrase-ological meanings:

- (12a) Egyptian Arabic *yišidd ša'ru* “tear one's hair”
 (12b) Chinese *sī chě tóu fā* “draw/pull (out) the hair”
 (12c) Japanese *kami wo kakimushiru* “tear out the hair”
 (12d) Korean *meo-ri-teol-eul jui-eo tteut-da* “pull/pluck out the hairs”

In antiquity, the gesture of tearing the hair was a sign of mourning, of bewailing the dead (first mentioned in the Iliad, 18.23ff, when Achilles mourns the death of Patroclus). It can be supposed that most of the European idioms in (11) go back to this classical origin, even though the cultural dimension of the gesture, its semiotic meaning and anchoring

in the ancient culture, was lost in the course of history. Since the same idioms occur in languages that have little to do with the classical occidental culture, however, we should also take the possibility of independent formations into consideration. Similar gestures of grief and despair (causing oneself pain and actually tearing out one's hair) are widespread far beyond the Greek/Hellenistic sphere of influence. It is perfectly conceivable, therefore, that the corresponding idioms developed independently in different languages (cf. e.g. Levin-Steinmann 2004).

(ii) The morphosyntactic and lexical structures of these idioms do not raise particular problems. All the different verbs belong to the same semantic class (*tear, pull, pluck, draw, tear out, pull out*). Some languages (Dut, Pol, Fin and Trk) have an additional element "out off/from one's head". The minor languages North Frisian and Lëtzebuergesch do not provide an infinitive form but use the idiom in utterances such as "it's enough to tear one's hair out" (cf. Filatkina 2005: 206). Such pragmatic variants do not hinder us from counting them among the same idiom type.

(iii) The phraseological meanings of the idioms given by our informants vary between 'to be in deep despair, to show extreme desperation' and 'to be horrified by sth., to be frustrated or worried about sth.'; thus, they do not contradict the assumption that we are dealing with one widespread idiom type.

5. Outlook

The few examples have sufficiently demonstrated the existence of widespread idioms (i.e. idioms spread across approx. 40 languages), a fact that hitherto could not have been stated that clearly with respect to idioms, although it may have been clear for proverbs. We can assume that about 100-150 idioms reveal a very similar pattern of distribution. The bulk of these idioms consist of items that go back to well-known literary sources, above all to the Bible.

Out of the numerous questions resulting from the empirical material, only a few could be discussed in the present paper. We could observe how certain idioms are spread far beyond the European languages, while others show distinct gaps on the map. We then discussed the question to what extent deviations of the phraseological meanings can be

tolerated within the limits of one postulated “widespread idiom”. We have deliberately refrained from questions regarding the causes of wide distribution (common source, literary or contact-based borrowings) and pointed out that identical idioms may well have been spontaneous coinages in different languages. Further results were only touched upon: English is far from always representing the “giving” language, as is so often claimed. The examples analysed so far show that the various languages of northern Europe (Germanic, North-Finnic and Baltic) are not marginal figures but central to the “European uniformity”. Since a large number of languages in the east and southeast of Europe have not yet been included into the analysis and because there is currently no access to their idioms, it is only with reservations that we can speak of true “widespread idioms in Europe”.

I would like to express my thanks for help with the foreign idioms: to Eldrid Aas, Nils Århammar, Winfried Boeder, Aina Budvytyte, Robert Darquenne, Adile Ęmirova, Ken Farø, Natalia Filatkina, Ulrike Frenademez, Darina Genadieva, Eva Glenk, Ania Grzybowska, Manfred Gross, Erla Hallsteinsdóttir, Mohsen Hantout, Krisztina Hevér, Susanne Hose, Eun-Mi Hwang, Vida Jesenšek, Elinborg Joensen, Terje Keldoja, Reso Kiknadze, Sigrun Kotb, Varvara Leventoupoulo, Anna Litovkina, Flor Mena Martinez, Miriam Olejarová, Oleg Ostapovych, Antonio Pamies, Regula Schmidlin, Mudite Smiltena, Yanping Tan, Björn Wahlberg and Casia Zaharia.

References

- Abondolo, Daniel 1998 (ed): *The Uralic languages*. London: Routledge.
- Abquilina, Joseph 1972: *A comparative dictionary of Maltese proverbs*. Valletta/Malta: Royal University of Malta.
- Bächthold-Stäubli, Hanns 1933: *Handwörterbuch des deutschen Aberglaubens*, unter Mitwirkung von Eduard Hoffmann-Krayer, Bd. 5. Reprint 1987, Berlin & New York: Walter de Gruyter.
- Boeder, Winfried 1985: La structure du proverbe géorgien. In *Revue des études géorgiennes et caucasiennes* 1, 97–115.
- 1991: Struktur und Interpretation georgischer Sprichwörter aus Chewußuretien. In Sabban, Annette/Wirrer, Jan (eds): *Sprichwörter und Redensarten im interkulturellen Vergleich*. Opladen: Westdeutscher Verlag, 131–161.
 - 2005: The South Caucasian languages. In *Lingua* 115, 5–89.

- Braun, Peter/Krallmann, Dieter 1990: Inter-Phraseologismen in europäischen Sprachen. In Braun, Peter/Schaeder, Burkhard/Volmert, Johannes (eds). *Internationalismen. Studien zur interlingualen Lexikologie und Lexikographie*. Tübingen: Max Niemeyer, 74–86.
- Braunmüller, Kurt 1997: Kontaktlinguistische Probleme im Ostseeraum zur Zeit der Hanse. In Moelleken, Wolfgang, W./Weber, Peter J. (eds): *Neue Forschungsarbeiten zur Kontaktlinguistik*. Bonn: Dümmler, 81–88.
- 2004: Plädoyer für eine Umorientierung in der Beschreibung von Sprachwandelprozessen. In Nowele. *North-Western European Language Evolution* 44, 21–49.
- Burger, Harald/Dobrovolskij, Dmitrij O./Kühn, Peter/Norricks, Neal (eds) in print: *Phraseology. An International Handbook of Contemporary Research*. Berlin & New York: De Gruyter.
- Čermák, František 2004: Text Introducers of Proverbs and Other Idioms. In Földes/Wirrer (eds), 27–46.
- Comrie, Bernard 2005: Introduction to Caucasian. In *Lingua* 115, 1–4.
- Crystal, David 2000: *Language Death*. Cambridge: Cambridge University Press.
- Dahl, Östen/Koptjevskaja-Tamm, Maria 2001 (eds): *Circum-Baltic Languages*. Vol. 1: *Past and Present*. Vol. 2: *Grammar and Typology*. Amsterdam-Philadelphia: Benjamins.
- Dobrovolskij, Dmitrij O./Piirainen, Elisabeth (2005): *Figurative Language: Cross-cultural and Cross-linguistic Perspectives*. Amsterdam [etc.]: Elsevier.
- Filatkina, Natalia 2005: *Phraseologie des Luxemburgischen. Empirische Untersuchungen zu strukturellen, semantisch-pragmatischen und bildlichen Aspekten*. Heidelberg: Winter.
- Földes, Csaba/Wirrer, Jan 2004 (eds): *Phraseologismen als Gegenstand sprach- und kulturwissenschaftlicher Forschung. Akten der Europäischen Gesellschaft für Phraseologie (Europhras) und des Westfälischen Arbeitskreises „Phraseologie/Parömiologie“ (Loccum 2002)*. Baltmannsweiler: Schneider Verlag.
- Gamrekeli, N./Hoffmann, T./Kadagidse, N. 1973: *Deutsch-georgisch-russisches phraseologisches Wörterbuch*. Zweite durchgesehene und erw. Auflage. Tbilisi: Ganatleba.
- Haarmann, Harald 1993: *Die Sprachenwelt Europas. Geschichte und Zukunft der Sprachnationen zwischen Atlantik und Ural*. Frankfurt [etc.]: Campus-Verlag.
- Haarmann, Harald 2002: *Kleines Lexikon der Sprachen. Von Albanisch bis Zulu*. München: Verlag C.H. Beck.
- Hanko, Urve/Liiv, Gustav 1998: *Inglise-eesti idioomi-sõnaraamat*. Tallin: “Valgus”.
- Hewitt, George 2005: North West Caucasian. In *Lingua* 115, 5–89.
- Hose, Susanne 1996: *Serbški prislowny leksikon*. Bautzen: Lusatia.
- Korhonen, Jarmo 1991: Konvergenz und Divergenz deutscher und finnischer Phraseologie. Zugleich ein Beitrag zur Erläuterung der Verbreitung und Entlehnung von

- Idiomen. In Palm, Christine (ed) „EUROPHRAS 90“. *Akten der internationalen Tagung zur germanistischen Phraseologieforschung Aske/Schweden 12.–15. Juni 1990*. Uppsala: Acta Universitatis Upsaliensis, 123–133.
- Kuusi, Matti 1957: *Regen bei Sonnenschein. Zur Weltgeschichte einer Redensart*. Helsinki: Suomalainen Tiedekatemia.
- 1985: *Proverbia Septentrionalia. 900 Balto-Finnic proverb types with Russian, Baltic, German and Scandinavian parallels*. Helsinki: Suomalainen Tiedekatemia.
- Levin-Steinmann, Anke 2004: Lehnbildung oder Produkt der Wirkung vorbildsunabhängiger kognitiver Prozesse? In Palm-Meister (ed), 273–284.
- Mejri, Salah 2004: La conceptualisation phraséologique: une construction conceptuelle langagière ou la prégnance du concept par la langue. In Földes/Wirrer (eds), 59–73.
- Menac, Antica 1987: Gemeinsame semantische Gruppen in der Phraseologie der Europäischen Sprachen. In Burger, Harald/Zett, Robert (eds): *Aktuelle Probleme der Phraseologie. Symposium 27.-29. 9. 1984 in Zürich*. Bern [etc.]: Peter Lang, 269–289.
- Mieder, Wolfgang 2004: *Proverbs. A Handbook*. Westport & London: Greenwood.
- Mokienko, Valerij M. 1998: Phraseologisierung von Europäismen oder Europäisierung von Phraseologismen? Divergente und konvergente Prozesse in phraseologischen Systemen europäischer Sprachen. In Eismann, Wolfgang (ed): *Europhas 95. Europäische Phraseologie im Vergleich: Gemeinsames Erbe und kulturelle Vielfalt*. Bochum: Brockmeyer, 539–555.
- Naumann, Hans-Peter 1989: Typen niederdeutsch-nordischer Interferenz im Bereich der Phraseologie. In Hyldgaard-Jensen, Karl/Winge, Vibeke/Christensen, Birgit (eds): *Niederdeutsch in Skandinavien II. Akten des 2. nordischen Symposions ‚Niederdeutsch in Skandinavien‘ in Kopenhagen 18.–20. Mai 1987*. Berlin: Erich Schmidt Verlag, 241–259.
- O. Nagy Gábor 1979: *Mi fán terem? Magyar szólásmondások eredete*. (Where is it from? The Origins of Hungarian Common Sayings) [a harmadik, bovitett kiadást sajtó alá rendezte Kovalovszky Miklós]. Budapest: Gondolat.
- Őim, Asta 1993: *Sõnaraamatud – fraseoloogiasõnaraamat* [accessible on the Internet www.eki.ee].
- Paczolay, Gyula 1994: *European, Far-Eastern and some Asian Proverbs. A Comparison of European, Chinese, Korean, Japanese, Vietnamese and Other Asian Proverbs*. Veszprém: University of Veszprém.
- 1997: *European Proverbs in 55 Languages with Equivalents in Arabic, Persian, Sanskrit, Chinese and Japanese*. Veszprém: Veszprémi Nyomda Rt.
- Palm-Meister, Christine 2004 (ed): *EUROPHRAS 2000. Internationale Tagung zur Phraseologie vom 15.-18. Juni 2000 in Aske/Schweden*. Tübingen: Stauffenburg Verlag.

- Piirainen, Elisabeth 2000: *Phraseologie der westmünsterländischen Mundart*. Teil 1: *Semantische, kulturelle und pragmatische Aspekte dialektaler Phraseologismen*. Baltmannsweiler: Schneider Verlag.
- 2004a: Culture in figurative language: “Standard Average European” vs. dialects. In Baccouche, Taieb/Burger, Harald/Haecki-Buhofer, Annelies/Mejri, Salah (eds). *L’espace euro-méditerranéen: Une idiomaticité partagée. Actes du colloque international (Hammamet 19, 20 & 21 septembre 2003)*, vol. 2. Tunis: Ceres, 339–354.
 - 2004b: False friends in conventional figurative units. In Földes/Wirrer (eds), 157–167.
 - in print a: *Widespread Idioms: Cross-linguistic and Cross-cultural Approaches*. In Häcki Buhofer, Annelies & Burger, Harald (eds). *Phraseology in Motion. Proceedings of EuroPhras Basel 2004*. Baltmannsweiler: Schneider verlag.
 - in print b: Phrasemes from a Cultural Semiotic Perspective. In Burger et al. (eds).
 - in print c: Dialectal Phraseology – Linguistic Aspects. In Burger et al. (eds).
- Piñel, Rosa Maria 2004: Multilaterale Phraseologie (Spanisch, Deutsch, Arabisch, Französisch, Polnisch, Provenzalisch und Russisch). In Palm-Meister (ed), 363–372.
- Porsch, Peter 1997: Phraseologismen im interkulturellen Vergleich. In Barz, Irmhild/Schröder, Marianne (eds): *Nominationsforschung im Deutschen. Festschrift für Wolfgang Fleischer zum 75. Geburtstag*. Frankfurt/M. [etc.]: Peter Lang, 191–200.
- Price, Glanville (ed) 1998: *Encyclopedia of the languages of Europe*. Oxford: Blackwell Publishers.
- Stolz, Thomas (ed) 2001: *Minor Languages of Europe. A Series of Lectures at the University of Bremen, April-July 2000*. Bochum: Brockmeyer.
- van der Auwera, Johan (ed) 1998: *Adverbial Constructions in the Languages of Europe*. Berlin: Mouton de Gruyter.
- van den Berg, Helma 2005: The East Caucasian language family. In *Lingua* 115, 147–190.
- Yurtbasi, Metin 1993: *A Dictionary of Turkish Proverbs*. Ankara: Turkish Daily News.
- Yurtbasi, Metin 1996: *Turkish Proverbs and Their Equivalents in Fifteen Languages*. Istanbul: Serkon Etiket.
- Бушуй, Анатолий М. 1987: *Основные Вопросы Теории Фразеологии. Библиографический указатель по фразеологии. Выпуск 6*. Самарканд: СамГУ.
- Федотова, В. П. 2000: *Фразеологический Словарь Карельского Языка*. Петрозаводск: “Карелия”.

Abbreviations of European language names

Alb	Albanian
Blg	Bulgarian
Cr	Croatian
CrT	Crimean Tatar
Ctl	Catalan
Dan	Danish
Dut	Dutch
Eng	English
Est	Estonian
Far	Faroese
Fin	Finnish
Fr	French
Glc	Galician
Grg	Georgian
Grk	Greek
Grm	German
Hng	Hungarian
Ice	Icelandic
It	Italian
Lëtz	Lëtzebuergesch
Lith	Lithuanian
Ltv	Latvian
MrdE	Mordvin Erza
NFRs	North Frisian
Nor	Norwegian
Pol	Polish
Prt	Portuguese
Rmns	Romansh
Rom	Romanian
Rus	Russian
Slva	Slovakian
Slve	Slovene
Spn	Spanish
Swd	Swedisch
SwG	Swiss German
Trk	Turkish
Ukr	Ukrainian
Usrb	Upper-Sorbian

