

# Widespread Idioms in Europe and Beyond

*Toward a Lexicon of Common  
Figurative Units*

ELISABETH PIIRAINEN

This groundbreaking book in theoretical and empirical phraseology research looks at Europe's linguistic situation as a whole, including 74 European and 17 non-European languages. The occurrence of idioms that actually share the same lexical and semantic structure across a large number of languages has never been demonstrated so clearly before *Widespread Idioms in Europe and Beyond*. This book answers significant questions regarding hitherto vague ideas about the phraseological similarities between European languages and their cultural foundation. Starting from the intertextual origins of the idioms, the question which texts from ancient to modern times actually contributed to the "Lexicon of Common Figurative Units" now can be answered. The fact that once widespread motifs of folk narratives are among them is only one of the most striking and surprising results. This inventory, which analyzes 190 out of a total of 380 widespread idioms and includes maps, is valuable for academic teaching and further research in the fields of phraseology and figurative language, areal and contact linguistics, folklore, and European cultural studies.



Elisabeth Piirainen received her Ph.D. in German studies, Dutch philology, and Indo-European and general linguistics from the University of Münster, Germany. She has been a lecturer at the University of Jyväskylä, Finland, and a leader of projects on dialect research in Westphalia, Germany. Her intensive examination of a Low German dialect culminated in *Phraseologie der westmünsterländischen Mundart* (2000, 2 volumes), the first study on dialectal phraseology within a linguistic framework and recipient of the Johannes Sass Prize. This resulted in many further studies on phraseology, among them *Figurative Language: Cross-Cultural and Cross-Linguistic Perspectives* (2005, with Dmitrij Dobrovol'skij). Her publications have covered numerous different languages (including German, Dutch, English, Swedish, French, Russian, Greek, Finnish, Japanese) and a variety of fields, such as dialectal and areal phraseology, gender specifics, puns, the cultural foundation of conventional figurative language, and idioms in a European framework. Her knowledge of various languages, her familiarity with field research, as well as her long-standing expertise in phraseology research have allowed her to launch this large-scale project on widespread idioms.



# **Widespread Idioms in Europe and Beyond**



Alan Dundes  
*Founding Editor*

Wolfgang Mieder  
*General Editor*

Vol. 5



PETER LANG

New York • Washington, D.C./Baltimore • Bern  
Frankfurt • Berlin • Brussels • Vienna • Oxford

ELISABETH PIIRAINEN

# Widespread Idioms in Europe and Beyond

*Toward a Lexicon of Common  
Figurative Units*



PETER LANG

New York • Washington, D.C./Baltimore • Bern  
Frankfurt • Berlin • Brussels • Vienna • Oxford

**Library of Congress Cataloging-in-Publication Data**

Piirainen, Elisabeth.

Widespread idioms in Europe and beyond:  
toward a lexicon of common figurative units / Elisabeth Piirainen.  
p. cm. — (International folkloristics; v. 5)

Includes bibliographical references and index.

1. Idioms. 2. Figures of speech. 3. Language and languages—Globalization.
4. Language and languages—Foreign words and phrases.
5. Language and languages—Cross cultural studies.
6. Intertextuality. I. Title.

P301.5.I34P55 418—dc23 2012013834

ISBN 978-1-4331-0579-1 (hardcover)

ISBN 978-1-4539-0845-7 (e-book)

ISSN 1528-6533

Bibliographic information published by **Die Deutsche Nationalbibliothek**.

**Die Deutsche Nationalbibliothek** lists this publication in the “Deutsche Nationalbibliografie”; detailed bibliographic data is available on the Internet at <http://dnb.d-nb.de/>.

The paper in this book meets the guidelines for permanence and durability of the Committee on Production Guidelines for Book Longevity of the Council of Library Resources.



© 2012 Peter Lang Publishing, Inc., New York  
29 Broadway, 18th floor, New York, NY 10006  
[www.peterlang.com](http://www.peterlang.com)

All rights reserved.

Reprint or reproduction, even partially, in all forms such as microfilm, xerography, microfiche, microcard, and offset strictly prohibited.

Printed in Germany

FOR ILPO TAPANI





# TABLE OF CONTENTS

<b>Preface.....</b>	<b>1</b>
<b>List of Participants.....</b>	<b>5</b>
<b>1 Europe and Europe-wide Linguistic Studies.....</b>	<b>9</b>
1.1 The Concept of <i>Europe</i> : General Remarks and the Perspective of Widespread Idioms .....	9
1.1.1 <i>Europe: Name and Continent</i> .....	9
1.1.2 <i>Europe in the Context of Widespread Idioms</i> .....	11
1.2 Multilingual Approaches to Europe's Linguistic Diversity.....	13
1.2.1 <i>Early Europe-wide Linguistic Studies</i> .....	13
1.2.2 <i>Linguistic Typology</i> .....	14
1.2.3 <i>Eurolinguistics</i> .....	16
1.2.4 <i>Figurative Units in Europe-wide Linguistic Studies</i> .....	17
1.3 Figurative Language in a European Framework .....	19
1.3.1 <i>Metaphors</i> .....	19
1.3.2 <i>Paremiology</i> .....	20
1.3.3 <i>Idiom Research</i> .....	22
1.4 <i>Europeme, Euroversal, Europeanism or Internationalism?</i> .....	25
1.4.1 <i>Linguistic Typology</i> .....	25
1.4.2 <i>Lexicology</i> .....	26
1.4.3 <i>Phraseology</i> .....	27
<b>2 Conventional Figurative Language.....</b>	<b>31</b>
2.1 Idioms, Figurative One-word Units and Proverbs.....	31
2.1.1 <i>Figurative Units: Preliminary Remarks</i> .....	31
2.1.2 <i>Figurative Idioms</i> .....	33
2.1.3 <i>Figurative One-word Units</i> .....	34
2.1.4 <i>Figurative Proverbs</i> .....	38
2.1.5 <i>Subcategories of Widespread Idioms</i> .....	42
2.2 Cultural Foundation: Toward a Macro-structure of the Lexicon of Common Figurative Units.....	44
2.2.1 <i>The Problem of Grouping Figurative Units</i> .....	44
2.2.2 <i>Culture: A Constant Element of Figurative Units</i> .....	47
2.2.3 <i>Motivation of Figurative Units</i> .....	48
2.2.4 <i>Etymology of Figurative Units</i> .....	51
2.3 Intertextuality .....	54
2.3.1 <i>The Term Intertextuality</i> .....	54
2.3.2 <i>Manifestation of Intertextuality in Widespread Idioms</i> .....	55

<b>3</b>	<b>Widespread Idioms in Europe and Beyond:</b>	
	<b>Theoretical Approach</b> .....	59
3.1	Criteria and Working Definition of Widespread Idioms .....	59
3.1.1	<i>The Six Criteria</i> .....	59
3.1.2	<i>The Definition</i> .....	61
3.2	Explanation of the Criteria .....	62
3.2.1	<i>Geographical Distance</i> .....	62
3.2.2	<i>Genetic Distance</i> .....	63
3.2.3	<i>Lexical Structure and Image Component</i> .....	64
3.2.4	<i>Figurative Core Meaning</i> .....	66
3.2.5	<i>Historical Development</i> .....	67
3.2.6	<i>Origin and Cultural Foundation</i> .....	70
<b>4</b>	<b>Collection and Presentation of Widespread Idioms:</b>	
	<b>Empirical Approach</b> .....	73
4.1	Steps of Collecting the Linguistic Data .....	73
4.1.1	<i>A List of Potential Widespread Idioms and Preliminary Test</i> .....	73
4.1.2	<i>A Network of Competent Collaborators and the</i> <i>Questionnaires</i> .....	74
4.2	The Linguistic Situation in Europe .....	76
4.2.1	<i>The Indo-European Language Family</i> .....	76
4.2.2	<i>The Uralic Language Family</i> .....	78
4.2.3	<i>The Turkic Languages</i> .....	79
4.2.4	<i>The Caucasian Languages</i> .....	80
4.2.5	<i>The Semitic Language Family</i> .....	81
4.2.6	<i>Basque</i> .....	81
4.2.7	<i>Planned Languages</i> .....	81
4.3	Languages Represented in the Project .....	82
4.3.1	<i>Germanic</i> .....	82
4.3.2	<i>Celtic</i> .....	83
4.3.3	<i>Romance</i> .....	84
4.3.4	<i>Baltic and Slavonic</i> .....	85
4.3.5	<i>Albanian, Greek and Armenian</i> .....	85
4.3.6	<i>Finno-Ugric</i> .....	86
4.3.7	<i>Turkic</i> .....	87
4.3.8	<i>Georgian, Maltese, Basque and Esperanto</i> .....	87
4.3.9	<i>Languages Outside of Europe</i> .....	88
4.4	Presentation of the Data .....	89
4.4.1	<i>Detailed Version and Space-serving Version</i> .....	89
4.4.2	<i>Headings, Semantic Explications, Cultural Foundations,</i> <i>and Comments</i> .....	90
4.4.3	<i>The Data Block</i> .....	92

4.4.4	<i>Variants and Standard Dictionary Form</i> .....	96
4.4.5	<i>Translation Problems</i> .....	98
4.4.6	<i>Writing Systems</i> .....	99
4.5	Geographical and Sociolinguistic Aspects .....	100
4.5.1	<i>Areal Dimension: Maps and Grids</i> .....	100
4.5.2	<i>Lesser-used and Standard Languages</i> .....	102
<b>5</b>	<b>Antiquity as a Source of Widespread Idioms</b> .....	105
5.1	Widespread Idioms and Antiquity: Introduction .....	105
5.2	Classical Mythology and Narratives (A 1 – A 13) .....	107
5.3	Proverbial Phrases Current in Antiquity (A 14 – A 24) .....	129
5.4	Quotations from Greek and Roman Authors (A 25 – A 37).....	156
5.5	Antiquity as a Source of Widespread Idioms: Conclusion.....	168
<b>6</b>	<b>The Bible as a Source of Widespread Idioms</b> .....	171
6.1	Widespread Idioms and the Bible: Introduction .....	171
6.2	The Old Testament (B 1 – B 18) .....	172
6.3	Old and New Testament (B 19 – B 27) .....	199
6.4	The New Testament (B 28 – B 42).....	222
6.5	Concluding Remarks on the Bible as a Source of Widespread Idioms .....	251
<b>7</b>	<b>Various Ancient Sources of Widespread Idioms</b> .....	255
7.1	Widespread Idioms in Various Ancient Sources: Introduction .....	255
7.2	Antiquity, the Bible, and Other Texts from the Past (C 1 – C 11) .....	256
7.3	Blending of Ancient Textual Sources and Other Sources (C 12 – C 26) .....	284
7.4	Ancient Conceptions of the World (C 27 – C 36).....	322
7.5	Various Ancient Sources of Widespread Idioms: Conclusion .....	347
<b>8</b>	<b>Post-classical Literary Works as Sources of Widespread Idioms</b> .....	351
8.1	Widespread Idioms and Post-classical Literary Works: Introduction .....	351
8.2	Novels and Dramas of World Literature (D 1 – D 11).....	351
8.3	Literary Works of Modern Times, Light Fiction, and Cinema Films (D 12 – D 21) .....	374
8.4	Literary Works as Sources of Widespread Idioms: Additional Remarks.....	391

<b>9 Proverbs and Proverbial Units of Medieval and Reformation Times as Sources of Widespread Idioms .....</b>	<b>395</b>
9.1 Widespread Idioms and Proverbial Units of Medieval and Reformation Times: Introduction .....	395
9.2 Medieval Proverb Collections (E 1 – E 8).....	396
9.3 Medieval Proverb Illustrations (E 9 – E 14).....	415
9.4 Medieval Literary Texts (E 15 – E 23).....	427
9.5 Widespread Idioms Developed since Medieval and Reformation Times: Concluding Remarks .....	440
<b>10 Fables, Folk Narratives, and Legends as Sources of Widespread Idioms .....</b>	<b>445</b>
10.1 Widespread Idioms in Fables, Folk Narrations and Legends: Introduction .....	445
10.2 Fables of the Aesopic Type (F 1 – F 10).....	446
10.3 Jests and Comical Tales (F 11 – F 18) .....	469
10.4 Fairy Tales (F 19 – F 23).....	482
10.5 Animal Tales and Fabled Animals (F 24 – F 31).....	487
10.6 Concluding Remarks on Fables, Folk Narrations and Legends as Sources of Widespread Idioms.....	509
<b>11 Conclusions and Main Results .....</b>	<b>513</b>
11.1 Preliminary Remarks.....	513
11.2 Attempts at Explaining Cross-linguistic Similarities of Idioms.....	514
11.3 The Most Widespread Idioms and Probable Causes of Their Wide Spread .....	522
11.4 Sociolinguistic and Areal Linguistic Aspects .....	526
11.5 Outlook.....	530
<b>References .....</b>	<b>531</b>
<b>Abbreviations of the Language Names .....</b>	<b>561</b>
<b>Indexes .....</b>	<b>563</b>
List of Maps.....	563
Index of Literary Works .....	564
Index of Depictions .....	567
Alphabetical Listing of the English Widespread Idioms .....	567
Index of Further Linguistic Units .....	570
Index of Languages and Groups of Languages .....	577
Subject Index .....	580
Name Index .....	586

## PREFACE

The languages of Europe, with their intense cultural interconnections and their common intellectual and literary traditions, have been exposed to mutual influences for centuries, a fact that has led to a wide range of cross-linguistic similarities. These similarities have been observed at diverse linguistic levels, including the level of elements of the figurative language, such as idioms. The most intriguing question, however, namely which Europe-wide common idioms actually exist, has not yet been answered. This question was the first impulse for my book.

A good seven years ago, I sent an initial questionnaire that consisted of twelve idioms to about 40 idiom research experts of various languages – the result was encouraging: ten of these idioms showed lexical-semantic equivalents in almost all of the participating languages, ranging from Icelandic to Spanish, from Estonian and Latvian to Hungarian, Romanian, and Greek. The next task was to identify as many of such figurative units as possible by means of systematic investigation, that is, to discover idioms that exist in a large number of European languages (and even beyond) showing a similar lexical structure and figurative core meaning, for which I coined the term *widespread idioms*. After some preliminary tests, I came to the decision to include all European languages accessible to my research. My starting point was the idea that today, in an increasingly uniting Europe and with modern means of communication at our disposal, it should be possible, perhaps more easily than ever before, to find competent collaborators for many languages of Europe. Indeed, the number of languages involved in the project that eventually became “Widespread Idioms in Europe and Beyond” has grown to 74 European and 17 non-European languages.

With the help of about 250 competent collaborators who tirelessly filled in the many questionnaires for their native languages, we were able to identify a good 380 idioms that are truly widespread and thus form the core of a “Lexicon of Common Figurative Units”. Only half of these idioms can be discussed in detail in the present volume: For this volume, I selected those widespread idioms that can be traced back to identifiable textual sources, that is, that can be analyzed in terms of intertextuality.

The present book consists of four introductory chapters followed by a documentation section – a “Lexicon” in the sense of an inventory of widespread idioms and a reference work – comprising six comprehensive chapters, and a summary. The first chapter discusses Europe as a linguistic space and linguistic projects carried out across Europe. Chapter 2 deals with figurative language, its culture-based nature, motivation, and etymology of idioms, as

well as with the term *intertextuality* insofar as it is relevant for the present study. The third chapter looks at the theoretical framework of this book, develops criteria for distinguishing *widespread idioms* from related phenomena, and establishes a definition of the term. The fourth chapter describes the methodological approaches and the form of data presentation. It provides a good overview of the linguistic situation in Europe, including areal linguistic and sociolinguistic issues concerning the standard and lesser-used languages.

From the beginning, the goal has been to examine the widespread idioms in their cultural context and to clarify their origin. This challenging step made it possible to categorize the idioms according to their underlying text sources and the chronological layers that they may be assigned to, which led to the six-part structure of the documentation section. Chapters 5–7 deal with idioms that belong to the oldest layers of our “Lexicon”, which include texts from classical antiquity, the Bible, and various ancient sources. A smaller number of widespread idioms have their origins in post-classical works of literature; these are the topic of Chapter 8. Another post-classical layer of widespread idioms can be found in the form of proverb collections and anthologies of medieval and early modern times, as shown in Chapter 9. The tenth and last chapter of the documentation section analyzes idioms that are rooted in once widespread narrative motifs, in fables, tales or concepts of fabled animals.

Which idioms would actually be widespread and which would fall short of our criteria was rather unpredictable. Chapter 11 attempts to pursue the question of the causes of the idioms’ wide dissemination. Indeed, despite the seemingly heterogeneous manifestation of widespread idioms, some common tendencies and regularities can be detected: these are mostly different from what could have been expected, and there are surprising results.

Many people have contributed greatly to the development of this book, and I would like to express my heartfelt thanks to all of them. First and foremost, a most sincere note of gratitude should be given to all linguistic informants. It is only due to their native speaker idiom competence and their reliable, prudent, and careful work that this book can be published in the present form. Our informants evaluated extensive preliminary tests and a seemingly endless amount of questionnaires, they provided me with a wealth of valuable data, and they patiently answered many supplementary questions over the years.

No less important for the realization of this book were the people who helped me with contacting competent informants, especially for the small and lesser-used European languages. Speakers of the declining language varieties even checked the questionnaires within their family circles, together

with their parents or grandparents, which was often the only access to an idiom. Both groups of helpful collaborators, the informants and the contacts, are mentioned by name with great thanks in the section following this preface.

All participants were kind enough to work on a voluntary basis. Problems arose from some technical details which also were solved by individual readiness to help. From the beginning, the project “Widespread Idioms in Europe and Beyond” was supported morally and practically by the “European Society of Phraseology”, which was of great help to get the project under way. I am especially grateful to Dr. Natalia Filatkina and Dr. Ane Kleine-Engel who offered to affiliate the project with the Researchers Group “Historische Formelhafte Sprache und Traditionen des Formulierens” and the prestigious “Cultural Historical Research Center” at the University of Trier, Germany. This again meant a positive change for the project since now various scholars became aware of it and offered their willingness to cooperate.

Several informants who will be mentioned with thanks below have also made the effort to once again examine all idioms of their native languages. Still other people have taken the trouble to read the whole book or individual chapters for possible errors and further remarks or have contributed to it in other important ways. For their selfless assistance in checking the texts and suggesting improvements, my thanks go to Professor Dmitrij Dobrovol’skij, Roland Peiler M.A., and Professor Roumyana Petrova Stoyanova. I am especially grateful to Professor Wolfgang Mieder, who accompanied the book in all its stages of development with his encouraging words and constructive comments and who included it in the “International Folkloristics” Series.

It is almost impossible to summarize in a few words of thanks the wonderful collaboration of two further scholars who have both contributed fundamentally to this book with their rich philological and cultural knowledge and familiarity with numerous languages. Dr. Bettina Bock checked most instances of Greek and Roman authors cited in this book and found several inaccuracies. With her knowledge of the Slavonic and various other languages, she helped greatly to avoid a number of mistakes.

Dr. József Attila Balázs has rendered especially outstanding service to the book. He evaluated the entire manuscript and enriched it with his insightful remarks. As an expert of a large number of languages, including Korean and Chinese, he also noted inconsistencies in the writing systems used, in the literal translations of the idioms and in various other points. My heartfelt thanks go to him for this commitment of long duration. The responsibility for all remaining errors, however, is entirely mine.

This book, it is hoped, provides insight into far-reaching uniformities of idioms of the European languages. It should be seen as a first step toward an even more extensive study of figurative units common to many languages.

Elisabeth Piirainen

Steinfurt, Germany

March 2012



## PARTICIPANTS

I would like to express my greatest thanks to all linguistic experts and native speakers of the languages represented in this book who participated in the project “Widespread Idioms in Europe and Beyond”.

### 1 European Languages:

<b>Albanian</b>	Bujar Arifi, Ferika Berberi, Naxhi Feka, Driton Gashi, Vjosa Hamiti, Halit Jashari, Holte Mano, Annie Polatsek, Greta Qello, Milote Sadiku, Edisa Shahini, Ruzhdi Stringa
<b>Armenian</b>	Hasmik Ghazaryan, Anna Miansaryan, Ani Mkrтчhyan, Dora Sakayan
<b>Aromanian</b>	Joana Hadzi-Lega, Zoran Kostov
<b>Azerbaijani</b>	Mardan Aghayev, Yasemem Hasan, Joshgun Sirajzade
<b>Basque</b>	Koldo J. Garai, Iraide Ibarretxe Antuñano, Zuriñe Sanz Villar, Izaskun Uría Navaridas
<b>Belorussian</b>	Elena Berthemet, Nikolaj Kostjuk, Liudmila Stampfli-Labudzeva
<b>Bosnian</b>	Sanja Berberović, Nihada Delibegović
<b>Breton</b>	Gilles Boucherit, Francis Favereau, Erwan Vallerie
<b>Bulgarian</b>	Vladimir Barakov, Delyana Deneva, Ani Dimitrova, Darina Genadieva, Roumyana Petrova Stoyanova, Diana Stantcheva
<b>Catalan</b>	María Conca, Heike van Lawick, Sion Fernandez, Lladó Apollonia Nicolau, Javier Orduña, Francesc Pont, Ricard Wilshusen
<b>Cornish</b>	Wella Brown, Rod Lyon
<b>Croatian</b>	Maslina Ljubičić, Željka Matulina, Marija Omazić, Anita Pavić Pintarić, Lelija Sočanac
<b>Czech</b>	Hana Bergerová, František Čermák, Irena Pavlíčková
<b>Danish</b>	Ken Farø, Sonja Poulsen
<b>Dutch</b>	Lut Baten, Jef van Egmond, Kurt Feyaerts, Arend Quak, Karel F. Treebus, Dominique Van Peer, Truus de Wilde
<b>Erzya Mordvin</b>	Natalia Deviatkina
<b>Estonian</b>	Anneli Baran, Terje Keldoja
<b>Faroese</b>	Elinborg Joensen, Marius Staksberg, Detlef Wildraut
<b>Finnish</b>	Marjaana Moisio, Mira Nyholm, Paula Olsoni, Heikki Piirainen, Leila Tervonen

<b>Franco-Provençal</b>	(Aosta Valley) Saverio Favre
<b>French</b>	Martine Dalmas, Françoise Hammer, Gilbert Magnus, Ahmed Sadouki
<b>Galician</b>	María Álvarez de la Granja, Xesús Ferro Ruibal, Alba Vidal
<b>Georgian</b>	Ada Dadiani, Giorgi Mariamidze, Maja Pizchelaury, Inga Tsintskiladze
<b>Greek</b>	Marios Chrissou, Eleni Motsiou
<b>Hungarian</b>	József Attila Balázs, Csilla Timea Kacsó, Tamás Kispál, Tamás Pekáry, Roberta V. Rada
<b>Icelandic</b>	Erla Hallsteinsdóttir
<b>Irish</b>	Marcas Mac Coinnigh, Liam Mac Con Iomaire, Gearóid Mac Eoin, Liam Mac Mathúna, Máirín Nic Eoin, Caitríona Ó Dochartaigh
<b>Italian</b>	Grazia Biorci, Stefania Nuccorini, Sebastian Lucjan Matracki
<b>Karaim</b>	(West Karaim) Halina Kobeckaitė
<b>Karelian</b>	Ljudmila Dudka, Natalja Gorbels, Vira Shkolyarenko
<b>Kashubian</b>	Aneta Lica, Jerzy Treder
<b>Kazakh</b>	Gulnar Sarsikeyeva
<b>Komi-Zyrian</b>	Nikolay Kuznets
<b>Ladin</b>	Ulrike Frenademez, Veronica Irsara
<b>Latvian</b>	Anita Nacišcione, Mudite Smiltena
<b>Lithuanian</b>	Aina Būdvytytė-Gudienė, Saulius Lapinskas, Rūta Marcinkevičienė, Reda Toliekiene
<b>Low German</b>	Hein Brinkert, Christian Christiansen, Klärchen Denker, Anne Mentzer, Antonius Tervoren
<b>Luxembourgish</b>	Robert Clee, Natalia A. Filatkina, Anne Noesen, Fabien Olsem, Martine Schadeck, Melanie Wagner
<b>Macedonian</b>	Iva Fidancheva, Joana Hadzi-Lega
<b>Maltese</b>	Joseph M. Brincat
<b>Mari</b>	Ljudmila Barzeva, Lidia Matrosova, Sonja Tshesnokova
<b>Moksha Mordvin</b>	Juri Jufkin, Osip E. Poljakov
<b>North Frisian</b>	Nils Århammar, Friede Faltings, Julius Nickelsen, Adeline Petersen, Anna Schröder, Willy Schröder, Ingke Wögens
<b>Norwegian</b>	Eldrid Hågård Aas, Svein Viken Opsahl
<b>Occitan</b>	(Torino) Monica Cini, Ferrier Serafino; (Alta valle di Susa) Monica Cini, Renato Sibille

<b>Polish</b>	Paweł Bąk, Anna Gondek, Martina Jagielski, Katarzyna Kozak, Marek Laskowski, Sebastian Lucjan Matracki, Renata Szczepaniak, Anna Urban
<b>Portuguese</b>	Maria Celeste Augusto, José Luís de Azevedo do Campo, Guilhermina Jorge
<b>Provençal</b>	Raymonde Blanc, Albert Borel, Martine Dalmas, Nicole Verdon
<b>Romanian</b>	Anca-Diana Bibiri, Cornelia Guju, Maria Husarciuc, Casia Zaharia
<b>Romansh</b>	Gelgia Caviezel, Georges Darms
<b>Russian</b>	Natalia A. Filatkina, Sinaida Fomina, Kamilla Mamedova, Galina Mertsalova, Alexandr Neyburg, Yulia Polshiuk, Artëm Šarandin,
<b>Saami</b>	(Inari-Saami) Anna Idström, Hans Morottaja
<b>Sardinian</b>	Ignazio Putzu
<b>Scots</b>	Andy Eagle, Gavin Falconer
<b>Scottish Gaelic</b>	Tiber Falzett, John MacPherson, Effie Rankin
<b>Serbian</b>	Biljana Golubović, Manon Pejović, Zoran Velikić
<b>Slovak</b>	Peter Ďurčo, Monika Šajanková
<b>Slovene</b>	Kasilda Bedenk, Vida Jesenšek, Erika Kržišnik
<b>Sorbian</b>	(Upper Sorbian) Susanne Hose/Hozyna, Sonja Wölke
<b>Spanish</b>	Ramón Martí Solano, Carmen Mellado Blanco, Flor Mena Martinez, Silvia Molina
<b>Swedish</b>	Eva Danielsson, Gunnar Hammarskiöld, Mira Nyholm, Paula Olsson, Tommy Petersson, Mariann Skog-Södersved
<b>Swiss German</b>	Regula Schmidlin, Mirjam Weder
<b>Tatar</b>	Roza Ayoupova, Albina Kayumova
<b>Turkish</b>	Ayfer Aktaş, Didem Ozan, Zekiye Ölmez
<b>Udmurt</b>	Anna Baidullina, Aleksandr Yegorov
<b>Ukrainian</b>	Svitlana Melnyk, Oleg Ostapovych, Stefaniya Ptashnyk
<b>Venetian</b>	Giulia De Gasperi
<b>Veps</b>	Natalja Gorbel, Vira Shkolyarenko
<b>Welsh</b>	Angharad Dafis, Ian Hughes, Emyr Lewis, Malcom Williams
<b>West Frisian</b>	Frits J. van der Kuip, Willem Visser
<b>Yiddish</b>	József Attila Balázs, Ane Kleine-Engel

## 2 Non-European Languages:

<b>Akalanon</b>	(Akalan, Philippines) Melchor F. Cichon
<b>Arabian</b>	(Algeria) Ahmed Sadouki (Tunisia) Moufida Ghariani Baccouche (Egypt) Mona Noueshi, Sigrun Kotb (Morocco) Majda Filali
<b>Bété</b>	(Ivory Coast) Jean Philippe Zouogbo
<b>Chinese</b>	Lina Chen, Demin Kong, Yanping Tan
<b>Farsi</b>	Edward Ardeshtir Danesh
<b>Japanese</b>	Makoto Itoh, Taeko Nasu
<b>Khanti</b>	(North Khanty, Synja dialect) Sofya Onina
<b>Kirghiz</b>	Bolotbek Tagaev
<b>Korean</b>	Eun-Mi Hwang, Sujeong Jeong
<b>Mansi</b>	Svetlana Dinislamova, Katalin Sipőcz
<b>Mongolian</b>	Nyambayar Tsedenbal
<b>Telugu</b>	Chilukuri Bhuvaneswar
<b>Vietnamese</b>	Le Tuyet Nga
<b>Wolof</b>	Mame Couna Mbaye
<b>Yorùbá</b>	Olutoyin B. Jegede

## 3 Esperanto:

József Attila Balázs, Sabine Fiedler

## 4 The following individuals have helped with contacting experts of languages less accessible to idiom research:

(Budapest), Anneli Baran (Tartu); Stanisław Borawski (Zielona Góra); Bettina Bock (Jena); Attila Cserép (Debrecen); Martine Dalmas (Paris); Wolfgang Eismann (Graz), Peter Ernst (Vienna); Tiber Falzett (Edinburgh); Sabine Fiedler (Leipzig); Natalia Filatkina (Trier); Hermann Goltz (Halle); Anna Idström (Helsinki); Britta Juska-Bacher (Zurich); Terje Keldola (Tartu); Jarmo Korhonen (Helsinki); Saulius Lapinskas (Vilnius); Kathrin Müller (Munich); Astrid van Nahl (Bonn); Stefania Nuccorini (Rome); Gyula Paczolay (Veszprém), Roumyana Petrova Stoyanova (Rousse/Bulgaria); Annette Sabban (Hildesheim); Ingrid Schellbach-Kopra (Munich); Asmus Schröter (Helsinki); Vira Shkolyarenko (Sumy); Diana Stantcheva (Sofia); Hildegard L.C. Tristram (Freiburg).

# 1 EUROPE AND EUROPE-WIDE LINGUISTIC STUDIES

As its title suggests, the present book on *Widespread Idioms in Europe and Beyond: Toward a Lexicon of Common Figurative Units*, brings together two approaches to linguistics that have both produced rich literatures in their own right: studies on languages within a European framework, on the one hand, and studies on conventional figurative units, on the other. This chapter is devoted to views of Europe in so far as they are related – and likely to be beneficial – to our objective of studying idioms that are spread across a large number of European languages. Starting with general notes on the name and geographical area of Europe (Section 1.1), we will look at the conceptions of Europe in multi-language studies, especially in linguistic typology and areal linguistics, as well as in works on figurative language (Sections 1.2 and 1.3). This will be followed by a discussion of the terminology used in those linguistic works (Section 1.4). The latter three sections are, therefore, also to be understood as outlining the “state of the art” of the issues discussed in this book.

## 1.1 The Concept of *Europe*: General Remarks and the Perspective of Widespread Idioms

### 1.1.1 *Europe: Name and Continent*

The name of *Europe* is often associated with early Greek cults and Greek mythology, which have produced figures bearing this name.<sup>1</sup> Although a true etymological connection between the geographical name and the mythological figure called *Europa* is rather unlikely, these names were linked with each other even in Ancient Greek. The mythological *Europa* is well-known for the story of Zeus transforming himself into a tame bull and abducting her, the Phoenician princess, on his back, from Phoenicia/Asia to the island of Crete. Zeus then reveals his true identity and makes Europa the first queen of Crete. From ancient times and throughout history, this legend has found countless representations in poetry and art and lead to a strong symbol of Europe even in present days (cf. Dombrowski 1984; Steiner 1991).

The name *Europe* in a geographical sense occurs for the first time in the ancient Greek collection “Homeric Hymns”, which was composed by anonymous poets in the tradition of Homer, probably during the eighth through sixth centuries BC. The hymn in honor of the god Apollo belongs to the old-

---

<sup>1</sup> These figures are mentioned as *Εὐρώπη* or *Εὐρώπεια* in Hesiod’s “Theogonia” (357) and in Homer’s “Iliad” (14, 321), cf. Geisau (1967: 446f); Treidler (1967: 448f).

est, eighth-century layer;<sup>2</sup> and it is in this poem that *Europe* is mentioned in two passages (verses 251 and 290). The “Hymn to Apollo” places *Europe* in opposition to two other geographic concepts:

[H]ere I am minded to make a glorious temple, an oracle for men, and hither they will always bring perfect hecatombs, both those who live in rich *Peloponnesos* and those of *Europe* and all the *wave-washed isles*, coming to seek oracles.<sup>3</sup>

This context provides evidence for an approximate location for this early occurrence of *Europe* as a geographical term. According to the author, a large area, inhabited by many people, is covered by the fertile Peloponnese, Europe, and the islands surrounded by the sea (most probably the Aegean islands). We can conclude that the term *Europe* here refers to parts of Greece. In his modern translation of this passage, West (2003: 93) does not use the word *Europe* but uses “the Mainland”. Fuhrmann (2004: 21) sees a reference to the region east of the Aegean Sea, roughly corresponding to the European part of Turkey today, north of the Bosphorus, which was considered the boundary between Europe and Asia.

What followed was a varied history of the term *Europe*. Toward the end of antiquity, it can be found as a name for Macedonia or Thrace, among other things. Moreover, Herodotus’ tripartite division of the world into three continents: *Europe*, *Asia*, and *Libya* (i.e. *Africa*) was seen as binding throughout the ancient times (Treidler 1967: 449). During the following centuries, the concept of *Europe* extended more and more to the west. Only at the beginning of modern times does the name *Europe* prevail again, first as Europe of the nation states, and today, in a modern sense, preferably for a political and socio-cultural unit.

From a geological viewpoint, Europe can be seen as part of the Eurasian continent, given the common solid landmass of Asia and Europe, whereas political and cultural aspects usually predominate in the definition of Europe as a continent on its own. While this continent has clear northern, western, and southern boundaries in the form of coastlines, it is not clear how far it extends to the east. Since the 18<sup>th</sup> century the Ural Mountain Range has been regarded as its eastern boundary – with an extension to the Caspian Sea. However, this is an arbitrary border, based on neither geographical nor political nor cultural grounds. Its origins probably go back to old traditions of the fur trade between Russia and the colonial territory of Siberia. It was Vasily Nikitich Tatishchev (1686–1750) who in 1720 suggested that the borderline

<sup>2</sup> For a detailed discussion of this dating see Allen/Halliday/Sikes (1936: 183–186).

<sup>3</sup> We have chosen a traditional translation here (Evelyn-White 1967: 343). The Greek text reads: [...] ἡμὲν ὅσοι **Πελοπόννησον** πείραν ἔχουσιν ἡδ’ ὅσοι **Εὐρώπην** τε καὶ ἀμφιρῦτας κατὰ **νῆσους** [...] (“Homeric Hymns” 3, 251 and 290).

between Europe and Asia should be drawn along the Ural Mountains (cf. e.g. Bassin 1991; Stern 2009; Tornow 2010).

Europe is the second smallest continent in terms of geographical expansion (after Australia) and ranks second in terms of population density (after Asia). The various definitions of *Europe* largely depend on their objectives and intentions or the context in which they are used, highlighting either geographical, political, historical or cultural aspects. Any attempts to demarcate Europe from neighboring cultural and linguistic areas by means of distinctive features are doomed to failure. Culturally, it is impossible to establish clear boundaries. Rather, the roots of the so-called European occidental culture lie beyond the European continent, in the Middle East as well as in North Africa. Its early cultural origins extend even further east than the later Classical or Hellenistic world may suggest (cf. Fuhrmann 1995, 2004: 21–28; Munske/Kirkness 1996).

The linguistic situation is quite similar. The native languages estimated to be spoken in Europe amount to approximately 150 languages (or more than 200 according to other estimates, cf. Stolz 2001: 227), which are spread over six language families.<sup>4</sup> The major language family spoken in Europe, the Indo-European family, has substantially more representatives in Asia than it has in Europe. Moreover, Uralic and Turkic languages are spread on both sides of the Ural Mountains, and a large number of languages do not simply “stop” at any virtual border within the Eurasian continuum.<sup>5</sup> We will outline the geolinguistic situation of Europe and the involvement of individual languages in our project in Sections 4.2 and 4.3.

### 1.1.2 Europe in the Context of Widespread Idioms

As the book title *Widespread Idioms in Europe and Beyond* implies, our starting point is the languages of Europe. Our choice of the European continent as a whole derived from the idea that the complex linguistic situation in Europe might produce interesting results. Besides, there were practical reasons: Even a large-scale international project needs to restrict itself to a subset of languages, and the European languages offered a manageable group for this. The basis of our study is those languages of Europe that were accessible for our idiom research. This resulted, after numerous efforts, in 73 European languages which now take part in the project. From the beginning,

---

<sup>4</sup> By global comparison, Europe is not particularly rich in languages: According to Haarmann (2010: 112), it is the continent with the least number of languages – a situation which is likely to change over the next decades.

<sup>5</sup> For detailed discussions of this phenomenon see, among others, Haarmann 1993, 1995; Gamkrelidze/Ivanov 1995; van der Auwera 1998; König/Haspelmath 1999; Stolz 2001; Haspelmath 2001.

experts of non-European languages offered their participation, and our study also comprises 17 languages spoken in Asia and Africa, as well as Esperanto. As various examples have shown, there can be continua of widespread idioms, be it from Europe to the languages spoken in the Eurasian border area or to the Arabic dialects of Northern Africa. Therefore, the focus of the “Lexicon of Common Figurative Units” is not so much on “European”, and the addition *and Beyond* is an important component of the title of this book.

This study starts from the geographical definition of Europe – thus following the custom of most Europe-wide linguistic studies, despite the difficulties outlined above. Some of the languages considered here belong to both the European and the Asian continent (such as Turkish) or cannot be clearly assigned to one of them (e.g. Armenian, Georgian, cf. Bossong 2010: 371), while other languages spoken mainly in Europe have speakers beyond the boundary as well (e.g. Komi) or a language spoken mainly in Asia has also speakers in Europe (Kazakh). All this is no reason to exclude these languages from our study.

It would make little sense for the present linguistic study to start from a cultural definition of Europe based on extra-linguistic factors. There are several authors who think of Europe as a cultural area displaying various degrees of a “cultural Europeanness”, with a full-fledged center and a “lesser European” periphery, among other things.<sup>6</sup> The religions in Europe are sometimes regarded as an important component in such a context, with Christianity in its Western form in the central parts of Europe, the more distant Eastern Orthodoxy, and Islam and other religions at the periphery. However, there is no direct connection between culture – which manifests itself, for example, in the religious affiliation of the speakers of a language community – and the figurative language used by these speakers. The spread of cultural features and the spread of linguistic features are independent categories, and cultural and linguistic areas do not necessarily coincide. Thus, *Europe*, as used in this book, will be understood in terms of its geographical boundaries, extending eastwards as far as the Ural Mountain Range and including the Caucasus region.

In the sense of a national romanticism, earlier phraseology research used to focus on the idea that idioms were unique to a particular language in question, like a mirror of the national culture or a national mentality, an idea that has been disproved since. We must now avoid a new romanticism of trying

---

<sup>6</sup> Such views can be found, among other things, within “cultural-anthropological” approaches to the linguistic situation of Europe where cultural conditions and the linguistic situation are confused. The exclusion of the “Orthodox world” or the Islamic states in Southeastern Europe from Europe is based on the biased assumption that cultural and linguistic features develop in parallel ways. See Voss (2010) for further discussion.



to draw a dividing line between the languages of Europe and those of other continents in an attempt to establish a European “community feeling” or, even worse, trying to reconstruct some kind of “European mentality” or European “world view” by means of the “Lexicon of Common Figurative Units”. In this context, I would like to refer to König (2010: 432), who emphatically rejects the idea of constructing a “European world view” by means of European typology research, although similar ideas may be desirable for parts of Eurolinguistics (see Section 1.2.3 below).<sup>7</sup> Instead, we let ourselves be guided solely by our data when it comes to drawing conclusions on the cultural and areal-linguistic structure of Europe. It would be presumptuous to postulate any form of exclusivity of “European idioms” or to adjust studies to a search of so-called *Europeanisms* (cf. Section 1.4), i.e. idiom equivalents restricted to the languages of Europe. In order to prove such “exclusivity”, as it has sometimes been claimed,<sup>8</sup> all the remaining approximately 6,500 languages spoken around the world would have to be analyzed in view of their entire figurative lexicon.

## 1.2 Multilingual Approaches to Europe’s Linguistic Diversity

In this Section, we will consider several studies which fit in with our Europe-wide approach as far as analyzing a large number of European languages is concerned. Mention should be made of several specialized linguistic studies that use Europe’s geographic frontiers as starting points, as our study does. We will have a closer look at two substantial research projects that have studied lexical and structural issues, respectively, for a large number of languages from all over Europe and that have drawn conclusions on the areal distribution of specific linguistic phenomena based on the map of Europe. It would seem that *widespread idiom research* could benefit from these methodological approaches.

### 1.2.1 Early Europe-wide Linguistic Studies

The many languages spoken in Europe and their genetic and structural diversity have drawn the attention of numerous scholars for a long time. Along-

---

<sup>7</sup> Hinrichs (2010), an advocate of *Eurolinguistics*, came to the conclusion – albeit without studying widespread idioms himself – that there is a “common world view” (*gemeinsame Weltsicht*) manifesting itself in proverbs and phrasemes which are spread across many European languages: “Redewendungen wie *‘gegen den Strom schwimmen’* sind – bei leichter Varianz – in fast allen Sprachen Europas verbreitet und zeugen von einer gemeinsamen Konzeptualisierung der Wahrnehmung” [Phrases like *‘to swim against the stream’* are – with slight variation – common in almost all languages of Europe and show a common conceptualization of perception] (Hinrichs 2010: 943).

<sup>8</sup> Compare Eismann (2010: 718–721), among others, for this claim.

side the famous Indo-European studies and Historical Comparative Linguistics, which mainly focussed on the genetic relationships between languages, individual philologies developed from the 19<sup>th</sup> century onwards. At the beginning, these various (national) philologies worked separately and they were mostly interested in one particular language or language family. Their principal aim was to identify the characteristics of that individual language and, accordingly, to list separating factors, i.e. differences from other related or unrelated languages. Europe-wide cooperation, across the borders of nations and language boundaries, only developed much later.

The first really large-scale research project that looked at the full diversity of the European languages was the *Atlas Linguarum Europae* (ALE). Its origins date back to the 1960s,<sup>9</sup> and the frontiers that it works with are neither political nor linguistic, but merely geographical. The ALE is a joint European project intending to compare possibly *all* languages and dialects spoken in Europe. It contains the six language families of Europe – from Iceland to the Caucasus – with a total of 22 language groups or subfamilies, which, in turn, consist of 90 individual languages and dialects. This huge geolinguistic research project is aimed at interpreting the heterogeneous data collected from 2,631 localities and putting them onto symbol maps. Seven fascicles of the atlas have been published to date; they give insight into early cultural-historic layers of the European lexicon and Europe's cultural past (cf. Vierck 2002a, 2002b, 2010).

### 1.2.2 Linguistic Typology

Although, for a long time, the languages of Europe have been the best-studied languages in the world, it was only in the recent past that linguists recognized their great similarity in terms of syntactic, morphologic or phonologic structures, especially in contrast to non-European languages. It was realized that this remarkable uniformity cannot be ascribed to the genetic affiliation of the languages because it is shared even by genetically unrelated languages. Particularly inspiring was the concept of a *Sprachbund*<sup>10</sup> applied to languages of different families which show linguistic similarities – first discovered among the languages of the Balkans in the early 1930s (cf. Haarmann 1976b: 77–96; Stolz 2006a: 389).

Various scholars should be mentioned here who paid attention to the overall convergence of the European languages and established the European

<sup>9</sup> The year 1970 must be regarded as the actual founding date and Antonius Angelus Weijnen as the founder (cf. Weijnen 1983f).

<sup>10</sup> The term *Sprachbund* is partly equated with *linguistic area* and partly distinguished from it; see Stolz (2006b) for a comprehensive overview of issues of definitions and theory of these terms.

areal typology. One of the earliest works is Ernst Lewy's "Der Bau der europäischen Sprachen" (1942). Based on 18 sample languages (which he thought to be representative of a larger number of geographically neighboring languages), he distinguished five subareas in Europe. Lewy's fundamental ideas of looking at the languages of Europe collectively from an areal-typological perspective and identifying the areal composition of Europe were further developed, among others, by Haarmann (1976a, 1976b), who used a sample of 65 languages (excluding the Caucasus region) and presented a list of 16 features common to all European languages that ranged from phonemic-phonotactic to morphological and syntactic features.

Later, these questions were examined in a much larger context, in the first place through the project *Typology of Languages in Europe (EUROTYP)*, which is probably the largest international linguistic project that has ever existed.<sup>11</sup> One of the goals of this comprehensive project was, again, to find linguistic features that are common to all European languages – that is, to verify the hypothesis of *Europe as a linguistic area* – and to study these phenomena for as many European languages as possible, also in contrast to non-European languages. The number of sample languages now amounted to 150 or more languages. A sample of 23 European languages was kept constant across all studies and supplemented with various further languages, ranging from Saamic languages to Greek and from Galician to the Caucasian languages. The project has promoted linguistic typology and areal linguistics significantly, and it serves as the essential source of information on morpho-syntactic commonalities and differences between the languages of Europe until today.

Haspelmath (2001: 1493–1501) put forward a list of 12 common morpho-syntactic properties and came to identify certain contiguous subareas within Europe. With the help of so-called *isopleths* or *quantified isoglosses* (lines showing the geographical distribution of languages that share the same number of features, cf. Dahl 2001: 1458; Haspelmath 2001: 1505), common features of numerous languages were charted onto the map of Europe and certain typological areas emerged, among them a linguistic area in the center of the European continent which was called the *European Sprachbund* or *SAE area*. Explanations of how these typological common features developed were found in the far-reaching contacts between European communities over the centuries.

---

<sup>11</sup> The project brought together more than 120 researchers from 20 countries, under the leadership of Ekkehard König and assistance of Martin Haspelmath (cf. König 1996, 2010 for an overview). It has produced a large number of publications, all of them based on substantial linguistic typology research, cf. e.g. van der Auwera 1998; Kortmann 1997, 1998; König/Haspelmath 1999; Dahl 1990, 2001; Haspelmath 1998, 2001; Feuillet 1998, 1999; Haspelmath et al. 2001; Bosson/Comrie 1998–2003; Plank 2003, to name only a few.

Both research projects, “Atlas Linguarum Europae” and “Typology of Languages in Europe”, have demonstrated how the study of common or widespread features of the European languages can be based on an extremely wide variety of linguistic data. They will help our widespread idiom research, both in view of their methodological range and notion of Europe and their empirical approaches of collecting data and plotting them onto the map of Europe.

### 1.2.3 *Eurolinguistics*

Many other works would need to be mentioned here which have expanded our knowledge of Europe as a linguistic area, for example Hinrichs’ collection of articles (2004), which shows that the European languages are shifting continuously from a more synthetic linguistic type (Eastern Europe) toward a more analytic type (Western Europe), or the broad survey of four of the most prominent common features by Heine and Kuteva (2006). The authors demonstrate how the languages of Europe are becoming increasingly similar, and this unifying process is accelerating and affecting many European languages, including those of different families such as Basque and Finnish.

The two latter studies belong explicitly to the linguistic branch called *Eurolinguistics* that I would like to touch on briefly here. *Eurolinguistics* is considered a relatively new research discipline; the term *Eurolinguistics* (German *Eurolinguistik*) was first used by Norbert Reiter (1991: 111). The foundation of Eurolinguistics is usually associated with the year 1999, when the Eurolinguistic Circle of Mannheim (ELAMA) was constituted (cf. Haarmann 1999; Ureland 2003, 2005). Its origins are sometimes seen much earlier, starting not only from the Eurolinguistic symposia and publications (Reiter 1991, 1999) but directly from the Prague Linguistic Circle of the 1920s with its theoretical and methodological innovations, which were alternatives to the historical linguistics of that time (Ureland 2003).

Eurolinguistics is a promising research concept that – like the aforementioned areas of research – attempts to examine European languages across individual philologies and independently of their genetic relationship or geopolitical situation. The Eurocentric ideas sometimes connected with this (cf. Reiter 1999; Leuschner 2001), however, must be avoided. Various interlingual concepts and levels of language systems are considered to belong to the Eurolinguistics scope of research, such as:

[...] transference, integration and translation of common European structures as well as phonemes, morphemes, words, phrases, word formation patterns, phraseologisms, syntactic patterns etc. (Ureland 2003: 13)

The reference to *phraseologisms* in this quotation might indicate a connection between *Eurolinguistics* and our project “Widespread Idioms in Europe

and Beyond". However, questions of phraseology and other aspects of the figurative lexicon have, for the most part, been ignored in Eurolinguistics<sup>12</sup> and Europe-wide research projects in general. Similarly, van der Auwera (1998: 815f) provides a list of twelve linguistic features common to European languages, whose final item is called "lexical and phraseological similarities" (ibid. 816). Just exactly which studies are referred to as being concerned with *phraseological similarities*, however, and what is meant by *phraseological similarities* remains entirely unclear.

#### 1.2.4 Figurative Units in Europe-wide Linguistic Studies

References to phraseology and figurative language can only be found occasionally in some of the studies produced by the Europe-wide large-scale projects mentioned above. König and Haspelmath (1999: 115), for example, notice that idiomatic parallels which, in the end, are based on loan translations, even extend into the everyday colloquial language, as is assumed for the common formula *au revoir*:

Die idiomatischen Parallelen, die letztendlich auf Lehnübersetzungen beruhen, reichen sogar bis in die alltägliche Umgangssprache: französisch *au revoir* = deutsch *auf Wiedersehen* = polnisch *do widzenia* = russisch *do svidanija* = ungarisch *a viszontlátásra* = italienisch *arrivederci* = niederländisch *tot ziens*, usw. [...]. Diese lexikalischen und kulturellen Ähnlichkeiten reichen jedoch nicht, um einen europäischen Sprachbund zu rechtfertigen.<sup>13</sup>

A more subtly differentiated analysis of these units in individual languages shows that the languages do not simply converge more or less accidentally but that language policy activities may be behind it. German authorities, for example, were known to have promoted *auf Wiedersehen* since World War I in an attempt to eliminate any French elements from German, such as *Adieu*, which was the popular expression of goodbye at the time (cf. e.g. Schürmann 1994: 62). Besides, some articles deal with idioms, although rather incidentally. All of these idioms happen to be widespread, as our surveys have shown (cf. 4.1).

<sup>12</sup> The recent "Handbook of Eurolinguistics" (Hinrichs 2010), after all, has dedicated a chapter to some aspects of this subject. In his article on "Phraseological similarities of the languages of Europe", Eismann (2010) reports, among other things, on European proverb research and on our widespread idiom project, adopting some results which were available on the project homepage at that time.

<sup>13</sup> "The idiomatic parallels which are based on loan translations in the end, extend even into the everyday colloquial language: French *au revoir* = German *auf Wiedersehen* = Polish *do widzenia* = Russian *do svidanija* = Hungarian *a viszontlátásra* = Italian *arrivederci* = Dutch *tot ziens*, etc. [...]. However, these lexical and cultural similarities are not enough to justify a European Sprachbund".

In the article mentioned above, König and Haspelmath (1999: 120) analyze the German idiom *Du gehst mir auf die Nerven* in the context of the discussion of external vs. internal possessors in the European languages (cf. the equivalent German *\*Du gehst auf meine Nerven*, which does not exist). Equivalents of the idiom are known in at least 57 European languages, and this idiom would be a promising example to study these typological features on a Europe-wide scale (e.g. English *to get on someone's nerves*, Welsh *mynd ar nerfau rhywun*, Finnish *käydä jkn hermoille*, meaning literally “to go on someone's nerves” vs. German *jemandem auf die Nerven gehen* or Czech *jít někomu na nervy*, both literally “to go someone on the nerves”, etc. (cf. Piirainen 2011d: 232f).

Let us look at two further widespread idioms that have been commented on in the context of areal typology of the European languages:

If lexical similarities between the European languages are discussed – for instance neoclassical compounding [...] or idiomatic structure (e.g. *ivory tower/torre d'avorio/Elfenbeinturm*, *as poor as a church mouse/pauvre comme un rat d'église/arm wie eine Kirchenmaus*) – then the last several centuries are the appropriate time frame for explaining the historical links, but the basic syntactic structures common to SAE languages must be older. (Haspelmath 2001: 1506)

For the idiom *as poor as a church mouse* (known in more than 40 European languages, cf. (F 26) in Chapter 10) the given time frame can be applied, depending on what is meant by “the last several centuries”. The first English instance dates from 1659 (Tilley 1950 No. C382) while the German equivalent has been documented in the register of pre-reformation proverb collections (Seiler 1923: 144). The idiom is not known in current French. This may be ascribed to intra-linguistic reasons because there is a similar French expression with a different meaning: *un rat d'église* ‘a religious, pious person’.

The time of origin of the idiom *to live/be in an ivory tower* (with equivalents in ca. 35 European and several non-European languages, see (D 10) in Chapter 8) is certain; it is the year 1837, when Charles-Auguste Sainte-Beuve wrote the poem “Pensées d’Août”, in which he describes the isolated lifestyle of the novelist Alfred de Vigny and suggests he lived in a secluded *tour d’ivoire*.

These examples may serve to illustrate that the figurative lexicon of the languages spoken in Europe has until now received little attention in Europe-wide studies, including Eurolinguistics. This is paralleled by a lack of a “European view” within *phraseology research*: This discipline has a rich tradition of cross-linguistic (mainly bilingual) studies but, until recently, did not look at the linguistic situation of Europe as a whole. However, the fields of figurative language and Eurolinguistics could mutually supplement and enrich one another, even though each has ignored the other so far. This leads us to the next section.

### 1.3 Figurative Language in a European Framework

In this section we will have a look at some studies from the field of figurative language which also start from a variety of (European) languages, although there is nothing that comes close to such substantial Europe-wide research projects as the “Atlas Linguarum Europae” and the “Typology of Languages in Europe” outlined above. Generally, we have to distinguish between *metaphors* – used here in the sense of freely created figurative expressions such as poetic metaphors or individual metaphoric uses of words and phrases – and conventional figurative units, which are part of the lexicon. For the most part, these lexicalized figurative units can be subsumed under *phraseology* (see Chapter 2 for more details). Although it is not common practice in phraseology to consistently separate *proverb research* from *idiom research*, this distinction must be observed for the purposes of the present study because of the different states of research of the two disciplines.

#### 1.3.1 Metaphors

The extensive similarities between the languages of Europe in the area of metaphors have attracted the attention of linguists for a long time. Verifications of this idea by means of rich linguistic data, however, are very rare. Most works on this subject do not go beyond some five or six languages or one limited region of Europe. Studies that do not meet our overall view of Europe will be left out of consideration here. Mario Wandruszka’s book titled “Die europäische Sprachgemeinschaft” (1990) may serve as an example: the work is often quoted in the context of Europe’s linguistic uniformity although the languages analyzed (German, English, French, Italian, and Spanish) cannot be said to cover the actual “European linguistic community”.

One idea that is also often quoted and discussed within the European framework is Weinrich’s (1976) concept of an *Abendländische Bildfeldgemeinschaft*, which translates roughly as “occidental (or Western) image field community”. The author assumes that there is a consonance of image-fields across the individual languages of the Western world.<sup>14</sup> He distinguishes between the two conceptual levels of *Bildspenderbereich* (“image donor field”) and *Bildempfängerbereich* (“image recipient field”) which, interconnected, yield metaphors in a broad sense.<sup>15</sup> Analyzing metaphors

<sup>14</sup> “Es gibt eine Harmonie der Bildfelder zwischen den einzelnen abendländischen Sprachen. Das Abendland ist eine Bildfeldgemeinschaft”. Weinrich (1976: 287)

<sup>15</sup> Weinrich coined these terms, whose meanings are very similar to those of *source domain* and *target domain* used in the Lakoffian Conceptual Metaphor Theory, cf. Lakoff/Johnson 1980; Lakoff 1987, 1993. Following this theory, we should also mention studies here which emphasize the similarities of metaphors in general, at the very abstract level of con-

created by well-known poets of the past and present, Weinrich proves the fact that there is no translation problem with metaphors used by the Western image field community. He exemplifies this observation, among other things, by the image donor field ‘coin, coinage’ which, mapped onto image recipient fields like ‘language, linguistic expression’ lead to a variety of “financial metaphors” across languages, all of them easy to understand for speakers of the other languages (*to coin a word, coniare una parola, forger un mot, cuñar una palabra*, p. 287). His examples, however, indicate that Weinrich – although speaking of the “Abendland” or Western world, is primarily concerned with the languages of Western Europe, or to be precise, some of its main standard languages, which became similar through a long-time, common literary tradition.

For reasons of space, we will not list other works that were possibly inspired by Weinrich’s concept. Instead, we should mention one salient cultural-semantic approach to the languages of Europe, namely Jirina van Leeuwen-Turnovcová’s (1990) study on the usage of ‘right’ and ‘left’ in European languages. Through broad interdisciplinary research, including archeology, early and medieval history, Indo-European linguistics and ethnology, the author provides insight into conceptualizations of several semantic fields, symbolizations and culture-semantic developments, all of which are shared by a wide range of European languages and beyond.

### 1.3.2 Paremiology

Proverb research has been carried out on a European, if not worldwide, scale from the very beginning. Thus, paremiology is familiar with terms like *common European proverb* (cf. Mieder 1986, 1999a, 2004a). Multilingual proverb collections and dictionaries exist in large numbers; see the bibliography in Mieder (2004b: 266–269). They enjoyed great popularity in Europe since the time of Humanism and Reformation. One of the most important proverb anthologies of this early tradition is Hieronymus Megiserus’ “*Parœmiologia Polyglottos: hoc est Proverbia et Sententiae complurium linguarum [...]*”. Megiserus constantly expanded his first edition (1592) to, at last, a collection of proverbs and proverbial sayings in more than 13 languages; the final and best known edition dates from 1605. Various other scholars presented extensive polyglot collections of proverbs, though not in as many languages, cf. Walter K. Kelly’s “A Collection of the Proverbs of All (European) Nations” (1859), Baron Louis Benas’ study “On the Proverbs of European Nations” (1877/1878) or Matti Kuusi’s substantial work (e.g. Kuusi 1985).

---

ceptual metaphors (e.g. KNOWING IS SEEING, HAPPY IS UP). However, such findings are considered “universal”; they go beyond the scope of European languages and are not relevant to our discussion.



None of these anthologies and other works on formulaic expressions differentiates between *proverbs* and *proverbial phrases* or *sayings* (we would call them *idioms* in more modern linguistic terminology) and they even use quite different terms.<sup>16</sup> Along the same lines, a strict separation between proverbs and other proverbial expressions is not needed for paremiology in general. Proverb studies are not primarily part of linguistics but of a wide-ranging ethnological-folkloristic and cultural framework. That is why contemporary multilingual proverb dictionaries usually contain proverbs and proverbial phrases side by side.

Some of today's multilingual dictionaries refer explicitly to Europe, such as the three-volume "Concise Dictionary of European Proverbs" by Emanuel Strauss (1994), a rich source of proverbs from 60 European languages. Just as famous is Gyula Paczolay's "European Proverbs in 55 Languages" (1997), which goes far beyond the languages of Europe. Paczolay was able to demonstrate that 106 proverbial units, many of which are of classical origin, are really widespread across the languages of Europe and most of them have counterparts in non-European languages.

There is yet another work to be mentioned here that started from European paremiology and studies linguistic and cultural parallels of proverbial units within a global dimension: Matti Kuusi's important book "Regen bei Sonnenschein. Zur Weltgeschichte einer Redensart" (1957). From today's perspective, it is hard to imagine how it was possible at the time, without our modern media, to gather such comprehensive material for research from languages across the world. Kuusi succeeded in recording more than 3,000 variants of the saying connected with the natural phenomenon 'When it is raining in the sunshine', 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.<sup>17</sup> Röhrich (1987) gives an overview of this "geographic-historical method" – developed main-

---

<sup>16</sup> Besides *proverbia*, there are terms like *sententiae*, *loci communes*, *adagia* or *παροιμία*, *γνώμαι*, etc. As studies on classical and medieval formulaic phrases have shown (cf. Hallik 2007), the linguistically defined terms (*idiom*, *proverb* or *phraseme*) cannot be adequately applied to the analysis of historical stages of the languages. This problem may be set aside when considering pre-modern materials; see already Faselius (1859: viiif) for this discussion, as well as recently Gibson (2010: 5).

<sup>17</sup> In his introduction, Kuusi (1957: 5–26) describes how he sent circulars to all researchers and research institutes within his reach and corresponded with scientists and non-professionals all over the globe. He also tells the reader about problems trying to get into contact with representatives of "difficultly accessible" countries. He feared that several postal items had got lost. Finally, the various alphabets and transcriptions of all the languages involved posed a considerable problem, for which Kuusi again received international support.

ly within the Finnish paremiology and ethnology at the time of Matti Kuusi and his academic predecessors.

Their works have shown that it is possible to collect immense linguistic data through collaboration of a large group of researchers, using methods of field work research such as surveys and correspondence and analyzing and evaluating these data within a long-term perspective. Like the research projects *ALE* and *EUROTYP* (cf. 1.2), the prominent paremiological works discussed here encouraged me to launch the large-scale project on widespread idioms.

### *1.3.3 Idiom Research*

The linguistic discipline of phraseology research (with idiom research in its center) looks back on a more than one-hundred-year old tradition. The studies on phraseology carried out by Charles Bally (1905, 1909), especially his influential classification of fixed expressions in the context of his “stylistique”, are usually regarded as the beginning of modern linguistic research into phraseology. Indeed it was Charles Bally who pointed out extensive cross-linguistic similarities between the languages of Europe that were even more striking than their differences – a fact he tried to link to a “European mentality”:

Même pour un observateur superficiel, les langues modernes des pays dits « civilisés » offrent des ressemblances en nombre incalculable, et dans leur incessante évolution, ces langues, loin de se différencier, tendent à se rapprocher toujours davantage. La cause de ces rapprochements n’est pas difficile à trouver; elle réside dans les échanges multiples qui se produisent de peuple à peuple, dans le monde matériel et dans le domaine de la pensée. [...] Appelons ce fonds commun, faute de mieux, la *mentalité européenne*. (Bally 1909: 22f)

Whether Bally also had idioms in mind here, and which modern languages exactly he was referring to, is not explicitly clear from his remarks. It should be noted, however, how even more than one hundred years ago there was an awareness of the far-reaching similarities between the languages of Europe. Despite this awareness, next to nothing was known, for a long time, about the actual similarities that exist between idioms in the European languages. While proverb studies were carried out on a multi-language scale from the beginning (cf. the works cited above), there is no tradition of Europe-wide idiom research that would be comparable to the prosperous international cooperation in the field of proverb research, a gap that was not even recognized until recently.

Compared with paremiology or the large-scale projects outlined in Section 1.2, idiom research is in a sad state. Only two of the European phyla (namely Indo-European and Finno-Ugric, a subfamily of the Uralic languages) are

sufficiently represented in the literature. The majority of European languages have hardly ever been the object of idiom research, in clear contrast to paremiology and linguistic typology. 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, thus neglecting a large number of European languages.<sup>18</sup>

Then there are the observations of numerous phraseology researchers who discovered that there are many equivalent idioms in two or more languages, even in genetically unrelated languages. Often researchers have meticulously described the subtle cross-linguistic differences and created classifications detailing various degrees of equivalence (cf. e.g. Korhonen 2007). As valuable as such results may be for the theory of phraseology and many practical purposes, they have not contributed significantly to our overall understanding of the common figurative language in Europe. Neither the numerous studies on idioms of various individual languages nor the equally comprehensive work on contrastive comparisons of the idioms of two or more languages have so far been able to change this. They were not able to name the actual idioms that have equivalents in many European languages and thus to determine which idioms can be counted among the core inventory of a lexicon of European figurative units.

Nevertheless, there is one exception and that is the early and outstanding multilingual work by the Finnish researcher Oiva Johannes Tallgren-Tuulio, “Locutions figurées calquées et non calquées. Essai de classification pour une série de langues littéraires” (1932), which has received too little attention in idiom research. Using material from at least 14 languages, the author discovered that there are many equivalent figurative units in several languages, even in genetically unrelated languages.<sup>19</sup> Tallgren-Tuulio did not differentiate between figurative multiword and one-word lexical units, and he stressed that explanations for the similarities could only be found if the history of these units was included in the research. Thus he obtained a number of *Hel-lenisms* that found their way from their original Greek-Latin area into Arabic and current European languages.

---

<sup>18</sup> Until recently, much of idiom research has been concerned with a few standard languages, initially with Russian, German, French, Hungarian, Finnish, as well as Czech, Slovak, Polish, but fairly recently also including Romanian, Croatian, Serbian, Bulgarian, Slovene, Spanish, Catalan, Icelandic, and Danish, and, in its infancy, also Greek, Italian, Lithuanian, and Latvian. Only a few lesser-used or minority languages have been studied thoroughly with respect to their idioms, such as Upper Sorbian, Kashubian, and Inari Saami (see Section 4.3 for more details).

<sup>19</sup> The languages are: Arabic, Classical Greek, Danish, Dutch, English, Finnish, French, German, Hebrew, Italian, Latin, Modern Greek, Spanish, and Swedish (including Finland Swedish).

In this context, Hubschmid's multilingual study (1988) on figurative units containing a verb for 'to burn' should be mentioned. The author is the first to refer to a possible connection between linguistic areas ("Sprachbünde") and the spread of certain idioms across Europe and Eurasia. Unfortunately, his article has received little attention.

Other works also occasionally noticed common features of the idioms of some European languages but did not match Tallgren-Tuulio's or Hubschmid's variety of languages. The similarities of idioms and proverbs in the languages of the Balkans were noted relatively early. Mile N. Tomici (1983) shows equivalent figurative units across nine partly unrelated languages.<sup>20</sup> As an explanation, he points to the similar living conditions of the people in the Balkan region, which are assumed to have produced similar phraseological units independently of each other. Antica Menac (1987) presents an inventory of several dozens of presumably common European idioms drawn from six related languages: Croatian, Russian, German, English, French, and Italian. Her analysis reveals a number of near-equivalent idioms in these languages. These cross-linguistic similarities are exclusively attributed to loan translations.

Another comparative study of idioms (Korhonen 1991) again covers nine European languages, including German, Finnish, French, Italian, English, Swedish, Russian, Hungarian, and Estonian. The article is structured synchronically, and it focusses on cross-linguistic equivalents and differences of German and Finnish idioms (cf. the traditional cross-linguistic equivalent types labelled as full or total, partial and non-equivalents), whereas the idioms of the seven other languages serve as a control group. More carefully than Menac (1987), Korhonen refers to the origins of the idioms and a so-called "European cultural heritage" (cf. Section 11.2), so that in some cases common sources and ways of borrowing from one language to the next become visible.<sup>21</sup>

Manfred Görlach's "Dictionary of European Anglicisms" (2001) is a special case: the *anglicisms* documented here comprise not only single words but also some idioms. Although only 16 selected languages (representative of a larger number) were studied, the book makes clear which loan idioms were the most popular ones in many languages by the end of the 1990s. Görlach provides interesting information on the use of some Europe-wide idioms such as *round table* 'an assembly for discussion, esp. at a conference', loan

<sup>20</sup> The languages are: Albanian, Arabic, Bulgarian, Greek, Macedonian, Romanian, Russian, Serbo-Croatian, and Turkish. Compare also Thomai et al. (1999) for the numerous cross-linguistic similarities of figurative units in the Balkan languages.

<sup>21</sup> The results could well be complemented by diachronic 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.

translations of which were found in all of the 16 languages analyzed, while e.g. *golden handshake* ‘payment given on redundancy or early retirement’ only has loan equivalents in three languages out of 16 (cf. Section 11.2).

Other attempts to analyze idioms across several languages fall far behind the above-mentioned studies. Some research papers are rather quick to label idioms found in no more than four languages as *Europeanisms*, *internationalisms*, *inter-phraseologisms* or even *universals*. Such studies do not really belong in the present section about figurative units in a European or multilingual context.<sup>22</sup> However, we will briefly look at them in the next section because of the terminology that they use.

### 1.4 *Europeme, Euroversal, Europeanism or Internationalism?*

In the history of linguistics, we find several attempts to capture the common characteristics of European languages by means of an adequate terminology. The most common term to refer to certain common properties found in numerous European languages is *Europeanism*. In typology, related terms such as *SAE-feature*, *Europeme* and *Euroversal* are used as well, while lexicology and phraseology prefer – besides *Europeanism* – the term *internationalism* and related terms. In this section, we would like to take another look at some of the studies mentioned above, but now from the viewpoint of *terminology*. We will concentrate on examples from *linguistic typology*, *lexicology*, and *phraseology* and consider whether we may benefit from the terms used in these disciplines for the purposes of our investigation into Europe-wide idioms.

#### 1.4.1 *Linguistic Typology*

It was in 1939 when Benjamin Lee Whorf coined the term *Standard Average European* (SAE) for a group of (mainly Western) Indo-European languages that shared many grammatical structures, using the term in contrast to languages with other types of grammar (cf. Whorf 1956: 138). The origin of this term must be seen in its historical context; Whorf coined it when he compared structures of Native American languages with European languages. The term met with criticism much later because of ideological, hegemonic connotations. Whorf used it, among other things, to describe a way of thinking of Europeans as opposed to that of the Hopi people, and not primarily as a term for common grammatical properties. Typology research, on the other hand, embraced this term, and it became generally established through the project “Typology of Languages in Europe” (cf. Section 1.2), where it is

<sup>22</sup> Other studies on European phraseology which are not centered on idioms can be disregarded here, e.g. Gréciano 1998, 2002a, 2004.

used in the sense of terms like *European Sprachbund* or *European linguistic area*. Individual pan-European commonalities were then sometimes referred to as *SAE-features*.

Terms closely related to *SAE-feature* include *Europeme*, *Euroversal*, and *Europeanism*. These terms are basically employed in works on linguistic typology and related studies. It would seem that they are used, for the most part, synonymously or interchangeably, while individual authors have their preferences. Van der Auwera (1998: 815f) summarizes a set of twelve linguistic features common to European languages, which he calls *SAE-features*. Several authors have followed this terminology: Haspelmath (2001) speaks of *major SAE-features*, while earlier works (Haarmann 1976a, 1976b) preferred the term *Europeme*.

The German expression *Europäische Universalien* (Haarmann 1976a: 105–153) sounds like an oxymoron at first (assuming that something could either be *European* or *universal*), but “Universalien” is used metaphorically here in the sense of ‘broadly spread across the languages’. The same can be said of the term *Euroversal* (a word amalgam of *Europe/European* and *universal*) for properties shared by a large number of European languages. This term is largely used by linguistic typologists describing common grammatical features of European languages (e.g. Kortmann 1998; Haspelmath 2001). Later studies have been more cautious in using these terms because languages outside of Europe may reveal the same features. As a curiosity, Heine and Kuteva (2006: 11, footnote 9) point to the example of Swahili, an African language that shares nine out of twelve Europemes and, therefore, would actually have to be considered “a European language”.

We do not want to go deeper into the criticism these terms have met with, be it because they have not been defined clearly or because they are not free of Eurocentric connotations, etc. Instead we can say, in summary, that these terms are unsuitable for our work. This is primarily due to the observation that it would be counterproductive to use the terms given that various idioms under discussion here have spread far beyond Europe.

#### 1.4.2 Lexicology

Some of the terms to be discussed here are also well-established in disciplines other than typology. The term *Europeanism* is especially well-known in the field of vocabulary and loanword research. There is a large literature on this topic in German, compare, among others, Braun et al. (1990) or Bergmann (1995) for an overview of the history and the various attempts to define this term. In most of these studies *Europeanism* refers to a wide range of lexical borrowings, loan translations, calques, and borrowed word meanings that can be classified as *Europeanism*, independently of the genetic af-

filiation of the languages to which these words belong (cf. e.g. Reichmann 1993). A large group of *Europeanisms* consists of lexical units that were adopted from Latin over two thousand years of history (cf. also the term *Euro-Latin*, Munske 1996; Munske/Kirkness 1996; Volmert 2004). We can often recognize a Eurocentric perspective here as many of these so-called *Europeanisms* have actually been studied only for a few major European languages.

Besides *Europeanism*, the term *internationalism* is also used in lexicology. It was originally used in works on technical language and translation studies. According to Hengst's definition (1978: 467), an *internationalism* is a loanword spread across several languages with phonetic and morphologic adaptation and a constant meaning. Consequently, the term (*lexical*) *internationalism* (or sometimes *interlexeme*) is known for words with an analogous formal and semantic structure in several languages. The rich literature on this topic has not succeeded in producing a uniform definition. A frequently discussed issue is the number of languages required to validate the status of internationalism. Several lexicologists argue that internationalisms should be represented in at least three languages which belong to at least two different language families (cf. e.g. Schaefer 1990).<sup>23</sup> A Eurocentric view has again been criticized because studies on internationalisms have almost completely ignored non-European languages (Matta 2002/2003).

*Europeanism* and *internationalism* are rather pretentious terms. A minimum number of three (or a few more) languages required by a definition, however, even makes them almost semantically empty. In view of our widespread idiom research, it soon becomes clear that we do not need to follow the terms discussed here.

### 1.4.3 Phraseology

Let us now turn to works which are concerned with cross-linguistic similarities of idioms and other figurative units (Section 1.3). They also use some of the terms discussed above but look back on a different linguistic tradition.<sup>24</sup> When classifying phrasemes according to their origins, we often find a dichotomy: first, there are those phrasemes that go back to ancient sources like antiquity or the Bible and are considered calques or borrowings in the language in question, and secondly those phrasemes that are considered "unique" to this particular language, drawn from internal sources and developed by their own speakers. For the first group, the term *internationalism* is well established in phraseology. Its terminological counterpart for the second

<sup>23</sup> See Bergmann (1995: 260–262) and Hausmann/Seibicke (1991) for further discussion.

<sup>24</sup> Another term, German *Europhrase*, was coined within the Indo-European Studies, cf. Bock 2001, 2009. We do not want to go into this term in more detail here.

group cannot be captured as clearly. In older works we find the term *nationalism*; other studies describe this concept in terms such as “the traditional domestic phraseology” (e.g. Ďurčo 2007: 732).

In his comparison of Russian and German phraseology, Rajchštejn (1980: 36f) uses the term *internationalism*. He probably is not the first one to use it, but his work had a great influence on later phraseology. Rajchštejn discusses different types of structural-semantic equivalence. He starts with the group of *FE-internatsionalizmy* (ФЕ-интернационализмы, “phraseological units-internationalisms”): These are “full equivalents” in Russian and German and may have equivalents in other languages. He contrasts them to phrasemes with a national specific or marker, such as a national name (e.g. German *nach Adam Riese*) which have no equivalents in other languages (so-called zero-equivalence). This dichotomy between “international” and “national” cannot be found as distinctively in works on lexicology using the term *internationalism*, but it runs through the history of phraseology and has contributed to coining a different meaning of *internationalism*.

Rajchštejn (1980: 36f) gives 13 examples of “internationalisms” found in Russian and German. Here, we will present the German idioms and add an English translation:

1	<i>den Augiasstall reinigen/ausfegen</i>	“to clean/sweep the Augean stable”
2	<i>zwischen Szylla und Charybdys</i>	“between Scylla and Charybdis”
3	<i>den Pegasus satteln/besteigen</i>	“to saddle the/climb on the Pegasus”
4	<i>der Gordische Knoten</i>	“the Gordian knot”
5	<i>das Trojanische Pferd</i>	“the Trojan horse”
6	<i>Buridans Esel</i>	“Buridan’s donkey”
7	<i>im Evakostüm</i>	“in Eve’s costume”
8	<i>babylonische Verwirrung</i>	“Babylonian confusion”
9	<i>(um) kein Jota</i>	“not (by) a jot”
10	<i>sein Pfund vergraben</i>	“to bury one’s pounds”
11	<i>ihre Zahl ist Legionen</i>	“their number is legions”
12	<i>das goldene Kalb anbeten</i>	“to worship the golden calf”
13	<i>päpstlicher als der Papst selbst</i>	“more pontifical than the Pope himself”

Our research reveals that seven of these idioms are widespread. Idioms 1, 2, and 5 will be dealt with in Chapter 5 of this book, “Antiquity”, and idioms 7, 9, and 12 in Chapter 6 “The Bible”. Idiom 4, which is from the domain of “History” will be discussed elsewhere. A definition of *internationalism* is missing in Rajchštejn’s work. Thus, the reason remains unclear why expressions like 10, 11 or 13 should be counted among “internationalisms”,<sup>25</sup> probably it is the allusions to biblical, classical or clerical scenarios. Nevertheless, terms like *internationalism* or *international idiom* have since been

<sup>25</sup> Our pre-tests (cf. Section 4.1) showed that none of them is spread across more than some four or five languages.



adopted uncritically in a number of studies; they are used even for similar idioms in a few Slavonic languages (e.g. Stepanova 1998) and can be found for an incoherent group of phrasemes, sometimes in contrast to *internal* or *domestic* phrasemes.

Therefore we can say that phraseology research uses the terms *internationalism*, *international idiom* as well as *Europeanism* rather carelessly, since they have never been defined within a theoretic framework and are often applied to idioms accidentally observed to share a similar lexical and semantic structures in very few languages. As one example of several other studies of this kind, let us mention Braun's and Krallmann's article (1990) reporting on their identification of "international proverbial phrases", which they claim were unknown at the time of publication. The authors use the German term *Inter-Phraseologismus* for this phenomenon and illuminate their "discovery" by means of long lists of quite similar idioms in German, English, French, and, in part, Italian. The question remains why precisely these languages should be considered representative of *internationality* – given the variety of languages spoken in Europe and worldwide.

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 (fables, the Bible, belles-lettres, or quotations from prominent people). Each idiom is examined for about three or four related European standard languages. Mokienko (1998) calls the tendency of these idioms to converge either *internationalism* or *Europeanism*, and in places even *universality*. According to the author (p. 543), a brief look into any recent idiom dictionary is sufficient to find an enormous amount of internationalisms.

All such listings (and these examples could easily be expanded) ultimately remain accidental; they must be seen as unmethodical and non-committal collections of heterogeneous material. Their main weaknesses are the small number of languages analyzed and a complete lack of theoretical foundation. It must be regarded as unacceptable to speak of *internationalisms* or *universals* without any systematic empirical data collection of – at least – languages from different continents and distant cultures. Thus, as terms in idiom research, *internationalism* or *Europeanism* are highly unsystematic and therefore unsuitable. What is missing is systematic empirical research, a working definition of the terms used, a catalogue of criteria by which idioms can be defined as *internationalisms/Europeanisms*, as well as a theoretical foundation of how to operationalize their compilation.

Therefore, we will refrain from using these terms and choose the neutral term *widespread idiom* (or *WI* for short) instead, for which a clear definition will be provided (Section 3.1).

## 2 CONVENTIONAL FIGURATIVE LANGUAGE

Since the aim of this study is to describe the “Lexicon of Common Figurative Units”, the concept of *figurative language* – along with *widespread idioms*, which will be clarified in the next chapter, and *Europe*, which we looked at in the previous chapter – is one of the key concepts in our work. In this chapter, we will turn to our research topic of prototypical idioms, or *figurative idioms*, and have a closer look at the term *idiom* and related terms.

Section 2.1 will briefly discuss the terminology of *figurative language* and *phraseology*. The distinctions between *conventional figurative units* and other types of metaphorical expressions will be fundamental to the present book. Another important area is the cultural foundation of figurative lexical units, including idioms. Idioms are not only elements of the language but also parts of culture. In Section 2.2, we will look at the relationship between figurative units and aspects of culture, and especially at the etymologies and cultural background of widespread idioms, which will help us find a macro-structure for the lexicon on which we focus here. This leads us to a discussion of the concept of *intertextuality*, as this volume will be restricted to figurative units that have their origins in existing texts (Section 2.3).

### 2.1 Idioms, Figurative One-word Units and Proverbs

#### 2.1.1 Figurative Units: Preliminary Remarks

In our search for figurative lexical units that are spread across a great variety of languages, we have restricted ourselves, out of the wider area of *conventional figurative units*, to the category of *idioms*. This category provides ample and sufficient material for our study, given that approximately 5,000 or more well-known idioms are estimated to exist in each individual language. Other categories of figurative units – among them, most notably, figurative one-word units including figurative compounds and, in part, proverbs – would be worth being studied in their own right with respect to their potential multilingual circulation. We will discuss them in this chapter (but not include them in the lexicon later) in order to complement our main object of study, figurative idioms.

In what follows, we will approach the notion of *figurative unit of the lexicon* from various sides, looking at terms like *metaphor*, *phraseology*, *idiom*, and *proverb*, among other things. There are three essential dichotomies to be considered, and they are interrelated: firstly, the distinction between *conventionalized* and *freely created* figurative units; secondly, the distinction be-

tween *figurative* and *non-figurative* language; and thirdly, the distinction between *one-word units* and *multiword units of the figurative lexicon*.

According to a number of research traditions, phraseology is a secondary semiotic system, in which phrasemes are regarded as second-order signs, composed of pre-existent first-order signs. Although the terminology of European phraseology research has varied considerably, with terms like *phraseology* or *formulaic language* and individual elements called *phrasemes*, *phrasal lexical items*, *fixed expressions* or *multiword lexical units* counting among the most common, there is a consensus in the discipline that *phrasemes* (or *phraseologisms* in older terminology) are conventional multiword expressions. There is a long tradition in European linguistics of classifying phrasemes into different classes such as *idioms*, *proverbs*, *restricted collocations* or *routine formulae*. The consensus here is that the central group of phrasemes is that of *idioms* (cf. Burger et al. 2007: 12). The most crucial property of idioms is their semantic irregularity, or idiomaticity.

Modern linguistics tends to subsume all manifestations of *formulaic language* under the term “phraseology”. While in principle, such a wide conception of phraseology may be an advantage, there are two phenomena that must be strictly kept apart: formulaic elements that belong to the language system (which are part of the mental lexicon), on the one hand, and all other kinds of stereotyped expressions, on the other.

Various phenomena of formulaicity belong to the latter group, ranging from formulaic text sequences such as all kinds of ritualized language, chants, prayers, nursery rhymes, narrative stereotypes (*Once upon a day*) to staged rituals, etc. While all of these belong to prefabricated stereotype language (cf. e.g. Stolz/Shannon 1976; Jarrett 1984), they are not elements of the language system. The same is true for formulae in poetic language, cf. the detailed study on Indo-European poetic text fragments involving myths and rituals by Gamkrelidze/Ivanov (1995: 731–740), and modern social rituals, formulaic small talk or formulaic sequences used by special occupational groups such as horse race commentators, weather forecasters, pump aerobics instructors, etc. (cf. Kuiper 2009). These are structurally fixed text segments, bound to particular kinds of texts or formulaic genres (cf. Wray 2002), but they do not classify as *phrasemes*. For historical phraseology, the border between the two groups can only be found by analyzing large text corpora (Filatkina 2009, 2010). Thus, *phraseology* is understood as the totality of fixed multiword units of a language, i.e. formulaic expressions that are elements of the lexicon and go beyond the level of a single word but do not go beyond sentence level.

### 2.1.2 Figurative Idioms

Since this book will concentrate on idioms proper, a comment on the term *idiom* seems appropriate. In English, *idiom* is sometimes used as an umbrella term for all types of fixed expressions, cf. Čermák (2004), who speaks of “proverbs and other idioms”. However, this usage is misleading because it is ambiguous and contradictory to the usage of terms like *Idiom* in German or *uduoma* in Russian. Our work focuses on prototypical idioms, i.e. idioms that have the prototypical property of *figurativity* (or *idiomaticity*, see below). The small group of non-figurative (or weakly figurative) idioms like *kith and kin* or *to and fro* can be disregarded here. All of our widespread idioms show a higher degree of figurativity.

Throughout this book, the term *idiom* will be used in the European tradition of phraseology research: The consensus here is that idioms are the central and most irregular group of *phrasemes*. In this tradition, *idiom* is defined by three constitutive characteristics: (i) *stability* (or *reproducibility*), (ii) *idiomaticity*, and (iii) *polylexicality*, a conception that most European phraseology researchers agree on today (cf. Burger et al. 2007).

(i) By *stability*, researchers usually understand the fact that idioms, as prefabricated units, are reproduced with approximately the same form and meaning, and are not constructed anew every time from individual words. Recently, however, corpus-based analyses have revealed that *stability* must be interpreted in a broader sense than had previously been assumed, allowing for certain variability. The notion of *stability* includes all kinds of variants of a given idiom in so far as they are *lexicalized*.

Therefore, we would give priority to the concepts of *lexicalization* or *conventionalization*. Idioms are elements of the language system; they are *conventionalized* in form and meaning and as such are part of the figurative lexicon. Because of this, they are clearly different from other word groups used in a figurative or metaphoric sense, namely freely created “novel” metaphors. Such figurative expressions are not retrievable from the mental lexicon but innovative coinages of the individual producer of a text.<sup>1</sup> The inclusion of a figurative expression in a dictionary may be considered a measure of its conventionalization, although printed versions usually lag behind actual usage.<sup>2</sup> Seen from the opposite perspective, idioms and other units of the

<sup>1</sup> Patrick Hanks (2006) coined the term *dynamic metaphors* for novel metaphors, in order to clearly distinguish them from *conventional metaphors*.

<sup>2</sup> For example, there is an entry in the “Penguin Dictionary of English Idioms” “*a banana skin* – a pitfall for the unwary which makes the victim look ridiculous” (PDEI 2001: 183). The German equivalent *eine Bananenschale* can be used in exactly the same figurative meaning, but there is no codification of this to be found in the paper versions of dictionaries so far.

figurative lexicon differ from other figurative expressions such as ad-hoc metaphors with respect to their degree of conventionalization (lexicalization).

(ii) *Idiomatcity* means that idioms (and other figurative units of the lexicon) are, in most cases, semantically irregular. Prototypical (or figurative) idioms can be interpreted on two different conceptual levels: on a primary level, i.e. their *literal meaning*, which underlies their inner form, and on a second level, their *lexicalized meaning*, also called the *figurative meaning*, which is the term I will use in this book. The so-called *image component* assumes the role of a semantic bridge between these two levels. The term *image component* refers neither to the etymology nor to the original image but to linguistically relevant traces of an image that are comprehensible for the majority of speakers. We are dealing here with an additional conceptual level, one that mediates between the literal reading (fixed in an idiom's lexical structure) and the figurative meaning of an idiom. It is precisely this characteristic of idiomatcity, or semantic ambiguity, which connects idioms with all other figurative units of the lexicon while separating them from that large portion of phrasemes which may fulfill criteria (i) and (iii) but which are not or are only weakly figurative. In this book, our research topic is not defined by the boundary between *multiword units* and *one-word units* of the lexicon but, independently of the number of lexemes involved, between *figurative*, i.e. semantically irregular units (such as figurative idioms, figurative compounds, conventionalized non-polylexical metaphors, etc.) and *non-figurative* units. Therefore, the theoretical framework of this book is the *Conventional Figurative Language Theory* (CFLT) outlined in Dobrovolskij/Piirainen (2005).

(iii) The study of idioms has been carried out, almost exclusively, in the discipline of phraseology, which regards *polylexicality* as a constitutive factor. Thus, idioms have been investigated together with phrasal verbs, light verb constructions, restricted collocations, and other fixed word combinations because they all consist of more than one word. However, the differentiation between *multiword* and *one-word lexical units* is not important to the study of figurative language, nor for producing a "Lexicon of Common Figurative Units". This leads us to the following section.

### 2.1.3 Figurative One-word Units

As has already been pointed out, conventional figurative one-word units are not listed in detail in this book because our object of research needs to be limited and idioms alone provide sufficient material. Nevertheless, it would

be desirable to investigate into one-word figurative units in the same way as into idioms and to study their distribution across languages.

In this context, we should distinguish mainly between two groups: firstly, lexical units – be they simple words or compounds – that can be used in a second, metaphorical or symbolic meaning, in addition to their primary meaning, and secondly, compounds where it is the combination of two lexical elements that can lead to a new, figurative meaning. The word *mole* may serve as an example of the first group. The name of the animal can be used metaphorically for an ‘underground agent’, based on several analogies between the literal and the figurative meaning: both the mole and the spy work ‘underground’, hidden from the view of others; both bring things to light by ‘digging them out’. According to the OED, this second, figurative meaning of *mole* was popularized by works of literature and has been in common use since the 1970s. Thus, the metaphoric meaning was conventionalized and has since then been a part of the lexicon. In the same way, a compound as a whole can be used metaphorically. For example, *springboard* is lexicalized in the figurative meaning ‘a good starting point for a career’ (motivated by the conceptual metaphor SUCCESSFUL IS UP), next to its literal meaning ‘springboard (e.g. at a pool)’.

Figurative compounds (compounds proper and juxtaposed compounds) are, naturally, found to a larger extent in synthetic than in analytic languages, due to the different word formation rules of these language types. Examples from English are therefore rare. As one of the few examples, *scapegoat* ‘an innocent person who is punished for the mistakes of others’ has been a topic of research. Like many idioms, it is of biblical origin and has undergone changes in the course of history. Pamies and Iñester (1999: 26f) have shown that equivalents of *scapegoat* in ten further languages reveal significant cross-linguistic differences due to different Bible quotations and translations.

The discussion of figurative one-word units, however, took place not so much in English but in German studies. We do not want to delve too far into the terminology suggested for this phenomenon such as *one-word phraseme* or *one-word idiom* (German *Einwortidiom*, cf. e.g. Henschel 1987) but will only consider the relevant points for the present study. As soon as we leave a more synthetic language (as German partly is), with its large amount of figurative compounds, and look out for equivalents in other languages, we are dealing with multiword units. They may have their origins in works of literature, ranging from Greek mythology (German *Zankapfel* “quarrel-apple” vs. “apple of discord” in other languages) or the Bible (e.g. German *Jammertal* vs. *vale of tears*) to Shakespeare’s works (German *Kanonenfutter* “cannon-fodder” as translation of *food for powder* in “Henry IV”) and more modern coinages like German *Salamitaktik*, from Hungarian *szalámitaktika*, vs. English *salami tactics*, *Geistesblitz* “mind-splash” vs. *flash of inspiration*,

*Schallmauer* “sound-wall” vs. *sound barrier*, *Regenbogenpresse* “rainbow-press” (cf. *yellow press*) to *Geldwäsche* “money-washing” vs. *money laundry*, etc.

Figurative one-word units have only occasionally been an object of research, in fact not for their own sake but in the context of *polylexicity* as one of the postulated defining criteria of phrasemes. The fact that a polylexical unit in one language may have a one-word equivalent in another language has been noted since the beginning of cross-linguistic studies. In her comparison of German and Hungarian phrasemes, Regina Hessky (1987: 62) refers to cases like German *unter vier Augen* “under four eyes” vs. Hungarian *négyszemközt* “four-eye-between”, both meaning figuratively ‘among the two of us, without the presence of other people’, or German *schwarz auf weiß* “black on white” vs. Hungarian *fehéren-feketén* “on white (paper) in black (letters)”, both meaning ‘reduced to writing; written down or in print’. At a complete loss of how to tackle this problem theoretically, researchers have even proposed to exclude such cases from cross-linguistic works, especially from bilingual phraseological dictionaries.<sup>3</sup>

We have encountered several such cases in our work on figurative units drawn from typologically different languages. They all are regular, following the morpho-syntactic rules of the given languages. Most often, such cases are figurative units with a compound structure in one language that correspond to idioms with a genitive or adjective-noun structure in another language. For example, Estonian *tõdehetk* “truth-moment” is a full equivalent of English *the moment of truth*, both figuratively meaning ‘a critical or decisive time, at which one is put to the ultimate test or important decisions are made’. In the same way, Estonian uses the compound *ajahamba* “time-tooth” as equivalent of *the tooth of time* and similar genitive constructions in most other languages (e.g. Finnish *ajan hammas* “the time’s tooth”), all meaning ‘the damages caused by time’.<sup>4</sup> It would seem odd not to consider such sets of figurative units to be full equivalents – even if they would be analyzed separately in traditional phraseology. The typological differences do not affect the im-

<sup>3</sup> The same problem has been observed from a monolingual perspective. By means of comprehensive Rhaeto-Romance data, Ricarda Liver (1989) has shown that it is impossible in that language to separate figurative compounds from polylexical figurative phrasemes. See also Korhonen (2007: 582) for this problem.

<sup>4</sup> Cf. the idioms (D 13) and (D 4) in Chapter 8 below. Icelandic, Hungarian, and Estonian are affected most by this type of word formation, cf. the widespread idioms discussed below: (A 3) *a/the Trojan Horse* vs. Icelandic *Trójuhestur*, (A 28) English *winged words* vs. Hungarian *szállóige* and Estonian *lendsõna(d)*, (B 2) German *im Adamskostüm* “in Adam’s suit” vs. Hungarian *ádámkosztümben* and Estonian *aadamaülikonnas*, (B 13) *a judgment of Solomon* or, with an adjective: German *ein Salomonisches Urteil*, vs. Icelandic *Salómons-dómur*.



age components and figurative meanings. This explains well why *conventional figurative unit* is much more suitable as a superordinate term for our study than the term *phraseme*.

Furthermore, we look at figurative one-word units here for yet another reason. Although they have so much in common with idioms, they have never been studied as extensively as idioms. Needless to say, they do not appear in idiom dictionaries, and they are partly overlooked in books on quotations or winged words as well. Possibly, including these units into the scope of research could enrich our knowledge of figurative language and its interrelation with culturally relevant texts.

Let us look at some examples from the Bible. The German language has a number of figurative one-word units of biblical origin, produced primarily by Martin Luther's bible translation. The German *Lippenbekenntnis* "lip-confession" is such a compound. It is lexicalized in its figurative meaning along the lines of 'a confession which is pretended, not really meant, manifesting itself only in words rather than deeds' (cf. English *lip service*). It does not require the co-occurrence of a particular verb or another lexical element; this may be the reason why it cannot be found in idiom dictionaries or appropriate reference books. The expression goes back to verses in the Old and New Testament in which people are warned against vain worship, repeating words only with their lips, without showing commitment to their meaning.<sup>5</sup>

Similar compounds of biblical origin include German *Schandfleck* "disgrace-spot" 'eye sore, something that spoils the aesthetic impression in an annoying way' (from Deut 32:5), *Erdenbürger* "earth citizen" 'human being' (from Ps 119:19), *Unschuldslamm* "innocence-lamb" 'sb. who is not to blame, not capable of doing evil' or *Hiobsbotschaft* "Job's message" 'message of great disaster' (cf. *Job's news*). Also belonging to these biblical units are figurative expressions like *Himmelfahrtskommando* "Ascension-commando" 'suicide mission', *Sabbatjahr* "Sabbath-year" 'a period of time where one is liberated from professional activity' (cf. *sabbatical year*) or *vorsintflutlich* 'very old and antiquated' (cf. *antediluvian*). Most of these expressions consist of two semantically autonomous elements which constitute the figurative meaning (e.g. *Lippenbekenntnis* 'a benefit performed only with the lips'); therefore, they are close to the idioms.

A special case is the German word *Feigenblatt* "fig leaf", also a compound. Its figurative meaning 'a means of concealing or veiling negative facts and events abashedly' cannot be derived from the elements FIG and LEAF; rather it is a clear intertextual allusion to the biblical story of Adam

<sup>5</sup> "[T]his people draw near me with their mouth, and with their lips do honour me, but have removed their heart far from me" (Isa 29:13); cf. also Matt 15:7 and Mark 7:6.

and Eve realizing that they were naked and feeling abashed.<sup>6</sup> It would be interesting to investigate in which of the European languages the figurative meaning of “fig leave” is lexicalized (such as in Dutch *vijgeblad*, French *la feuille de vigne*,<sup>7</sup> Russian *фиговый листок*, Hungarian *fügefalevél* or Finnish *viikunalehti*) and in which it is not. Idiom dictionaries have no information on this.

#### 2.1.4 Figurative Proverbs

Proverbs are not the main focus of this work. Paremiology has been practiced on a multilanguage scale from its very beginning. As outlined in Section 1.3.2, there exist comprehensive Europe-wide or worldwide proverb studies. In view of the question of the “wide spread” of proverbs, there was no research need as was the case for idiom research. Paczolay’s work (1997) alone deals with 106 proverbs and proverbial phrases that are spread across a large number of European and Asian languages. From the beginning of my investigation into “widespread idioms”, I decided not to touch on proverbs at all. In the course of my studies, however, it became evident that several widespread idioms are intertwined with proverbs. We will therefore take questions regarding proverbs into consideration, if only to the extent that our research project is concerned.

First, let us look again at the terminology, which is not fully consistent across phraseology and paremiology. Despite the extensive literature on proverbs, a generally acknowledged definition has not yet been arrived at (cf. e.g. Taylor 1931: 3; Whiting 1932; Kleiber 1989, 2005; Mieder 2004b: 2–4, 2007: 394–396. Phraseology usually regards proverbs as one class of *phrasemes*, since they show the defining characteristics of stability or conventionalization and polylexicity, and partly also figurativity. Indeed, by far not all proverbs are figurative. There are most common proverbs that differ from figurative units in so far as they can hardly be interpreted on two different cognitive levels. Compare, for example:

*Business before pleasure; Every country has its own customs; All’s well that ends well; To err is human; Everything has a beginning; There is an exception to every rule; Money isn’t everything in life; Nothing ventured, nothing gained; Opposites attract each other; No pains, no gains; A penny saved is a penny earned; Don’t*

<sup>6</sup> “[A]nd they knew that they were naked; and they sewed fig leaves together, and made themselves aprons” (Gen 3:7). The idea of the fig leaf is especially known by many iconographic representations in sculpture and painting.

<sup>7</sup> In French, the original expression *la feuille de figuier* “the leaf of fig tree” has later been replaced by *la feuille de vigne* “the leaf of vine”.

*make something out of nothing; Tastes differ; Never put off until tomorrow what you can do today; Work never hurts any man.*<sup>8</sup>

What these proverbs have in common is the fact that the literal readings coincide largely with the ideas expressed by these proverbs: no image component is needed to mediate between a literal and a figurative meaning. Thus, these proverbs belong to phraseology but not to figurative language. There are other factors at work that make them irregular units of the lexicon, for example archaic language elements or poetic features like rhyme, alliteration or assonance. However, there are certainly just as many obviously figurative proverbs, like *Make hay while the sun shines; No rose without thorns* or *Who keeps company with a wolf learns to howl*. For our study on “figurative” widespread idioms, only this latter group is of importance.

From the viewpoint of the linguistic discipline of *phraseology*, the features that distinguish proverbs from idioms have been described in detail in several places; for a survey, see e.g. Dobrovol'skij/Piirainen (2005: 49–53). These features are to be found on different linguistic levels, ranging from the formal side to the content plane, pragmatics, and the semiotic status of proverbs. To summarize briefly, all proverbs have a sentence structure; proverbs are general statements believed to express a universal truth; they often have a universal quantifier like *every, all, never*, etc.; many proverbs have the illocutionary force of recommendation using segments like *you don't, one must*; finally, proverbs have discursive autonomy: they are quoted as texts, having no deictic elements which would tie them to a current text or situation.

For the purposes of our book, we do not need to go into these properties in more detail nor discuss the various attempts at defining *proverb*. Instead, let us look at the concept *proverb* from the viewpoint of *paremiology*. Here, proverbs are seen as elements of verbal folk culture, together with other elements like *weather maxims, country sayings, riddles, counting-out rhymes, nursery rhymes*, and the like. From the paremiological perspective, *proverbs* and *proverbial phrases* (or *proverbial sayings*) are likewise elements of a code of cultural folklore. In order to study them, it is not important to differentiate between *proverb* and *proverbial phrase* by means of the linguistic properties mentioned above (i.e. because proverbial phrases do not share features of the proverb like sentence structure, conveying a general truth, showing an illocutionary force or discursive autonomy).

Nevertheless, we can also notice struggles for an adequate terminology. In his note on “Proverbial Phrases not Proverbs in Breughel's Painting”, Archer Taylor points to this issue, which his colleagues were less aware of at that time:

---

<sup>8</sup> These proverbs are taken from the “Dictionary of American Proverbs” (Mieder et al. 1992), with the first entry selected for each proverb.

The distinction between a proverb (*Der Apfel fällt nicht weit vom Stamm*) and the proverbial phrase (*To beat one's head against the wall*, *To bell the cat*) is old and has been perceived but rarely discussed by collectors of proverbs. It is not altogether clear, for example, just what distinctions Erasmus recognized. The "*Proverbia Communia*" cites exclusively or almost exclusively proverbs in the strictest sense. For this reason it is worth noting that Pieter Breughel (1525–1569) depicted only proverbial phrases (*sprichwörtliche Redensarten*) in his famous painting. The commentators [...] have not stressed this fact. (Taylor 1965)<sup>9</sup>

To summarize, in the area of "proverbial phrases", paremiology and linguistics (or phraseology) are concerned with the same *conventional figurative units*, although using different terms and dealing with rather different questions. The research topic of paremiology is proverbs (including proverbial expressions, proverbial comparisons, proverbial exaggerations, binary formulas, Wellerisms, etc.), while phraseology covers other kinds of phrasemes, among them idioms, for which the term *proverbial phrase* would not be appropriate. In this book, the term *idiom* is used, as discussed above.

The idioms analyzed in this book have connections with proverbs with respect to two areas: (i) several widespread idioms are derived from proverbs, and (ii), in some cases we cannot clearly determine whether the given figurative unit is an idiom using a sentence structure or rather (at the same time) a proverb.

Let us illustrate this with some examples. Several English idiom dictionaries include an expression such as *to look a gift horse in the mouth*, meaning 'to find fault with what has been given or be ungrateful for an opportunity'; cf. OID (1999: 151); CCID (2006: 154). The infinitive form and the text samples (*I'd be inclined not to look a gift horse in the mouth*, CIDI 2002: 154) all indicate that we deal with a normal verb idiom. In other languages, a truncated nominal phrase is common as well, e.g. French *un cheval donné* and Russian *дарёный конь*, literally "a given horse (given as a present)", or a one-word unit as in Finnish: *lahjahevonen* "gift-horse", all meaning figuratively 'a present with minute detriments (that one should not inquire of)'. We had included this expression in our questionnaires and according to our informants, it soon became clear that, unlike English, most languages do not possess the idiom versions but only equivalents of the full proverb "Don't look a gift horse in the mouth" as it has been handed down since Late Antiquity.<sup>10</sup> Accordingly, we did not continue our survey and removed the expression from the list of potential widespread idioms.

<sup>9</sup> Brueg(h)el and Erasmus actually dealt with both proverbs and proverbial phrases, but this is not the point here. Cf. also footnote 16 in Chapter 1.

<sup>10</sup> The first use of the proverb is attributed to St. Jerome (*Noli [...] ut vulgare est proverbium equi dentes inspicere donati* "Do not, as the common proverb says, inspect the teeth of a given horse", cf. Taylor 1931: 4; TPMA 9, 112–114) and its spread is ascribed to Erasmus' "Adagia" (4, 5, 24; Coll.W. 36, 156).

Frequently, however, the case is quite the reverse, i.e. the predecessor of a widespread idiom is an old proverb while in the current languages the proverb is not known anymore. Thus there are, for example, equivalents of the idiom *to foul/befoul one's own nest* in more than 20 languages, while the old proverb form "It is an ill bird that fouls its own nest" has been preserved in only a very few languages. Accordingly, we have included this idiom in our study (cf. (E 1) below). The origin of a proverb is not as clear for the equivalents of *to catch/clutch/grasp at a straw* (cf. (E 23) below) but can be reconstructed with the help of 16<sup>th</sup> century proverb collections. Chapter 9 deals with 23 widespread idioms originating from proverbs which emerged, for the first time, only in post-classical proverb collections or anthologies. The number of idioms found to originate from old proverbs is probably much larger. On the other hand, it is possible for one text source to have produced an idiom and a proverb in one and the same language, cf. *to cry wolf* and *Never cry wolf when not in danger*. Here it is Aesop's fable "The Boy who Cried Wolf" that lead to both figurative units; the idiom is not necessarily a truncated form of the proverb.

Most idioms have the syntactic function of sentence parts such as verb phrases, noun phrases, etc. Besides, there is a small group of *sentence idioms* (cf. 2.1.5): They have the structure of a sentence (like proverbs) but can clearly be separated from proverbs by means of other features. The widespread idiom *the die is cast* is such a sentence idiom. It is uttered in a specific situation, namely when an important decision about the future has been made and there is no turning back. Apart from its sentence character, the idiom is not at all similar to a proverb; what is lacking, in particular, is the factor of stating a "general truth" or an "apparent truth".

There are, however, some borderline cases since proverbs can also be used to comment on a very specific situation. For example, we could say *the devil is/lies in the detail* just while struggling with serious problems caused by apparently small, harmless things that are often overlooked. Therefore, this figurative unit could be classified as an idiom with the function of an utterance. Because it expresses a kind of folk wisdom or universal truth, however, it must be classified as a proverb and we excluded it from our collection of widespread idioms. The comment *speak/talk of the devil!* 'said when the person just mentioned appears unexpectedly' is a similar case (cf. Balázs/Piirainen 2013). Similarly, we are used to hearing *better late than never*<sup>11</sup> as a comment in a particular situation when something is finally being done although it should have been done earlier, or as an excuse for being

---

<sup>11</sup> This widespread expression is a literal translation of Latin *potius sero quam numquam*. It was used by various Roman writers and it became well-known mainly as a quote from Livy.

late. This utterance is closer to a proverb because of the universal quantifier *better – than*; besides it is only weakly figurative, and so we excluded it from our data as well.

### 2.1.5 Subcategories of Widespread Idioms

All of the 380 widespread figurative units we collected can be subsumed under the term *idiom* in a broad sense. These units belong to several subgroups of idioms, which will be briefly outlined here. The distinction between idioms with a sentence structure (*sentence idioms*) and idioms functioning as a part of a sentence (so-called *word idioms*) has already been mentioned above. Any subdivision of the latter group usually follows the grammatical categories of these “word equivalents” or their syntactic functions (i.e. *verb idioms*, *noun idioms*, *adverbial idioms*, and *adjective idioms*). Let us illustrate these and further subcategories by some examples, also in terms of their quantitative occurrence in the “Lexicon of Common Figurative Units”.

**Sentence idioms.** Our data include a small group of 15 widespread sentence idioms. Their main communicative function is commenting on a particular situation. Some of them show an explicit sentence structure, such as *the die is cast*; *the glass is half full*; *a time-bomb is ticking*; *all doors are open to someone*; *God (only) knows*, while others occur in an implicit sentence structure, e.g. *an eye for an eye*, *a tooth for a tooth*; *as if the earth had swallowed him up*; *after us/me the deluge*. These sentence idioms must not be confused with sentence-like predicative verb idioms such as the WIs *someone’s days are numbered*; *someone’s hands are tied*; *someone’s hair stands on end*.

**Verb idioms.** Most of the other widespread idioms have the communicative function of evaluating and characterizing someone or something. The ca. 280 verb idioms are by far the largest group among them, for example *to back the wrong horse*; *to tighten one’s belt*; *to swallow the/a bitter pill*; *to be on the same wavelength as someone*; *to rule with an iron hand*; *to follow the line of least resistance*; *to twiddle one’s thumbs*; *to vote with one’s feet*; *to be sitting on a powder keg*. This group also includes verb idioms with a more complex syntactic structure, e.g. *to know which way the wind blows*; *to sell the skin before you have caught the bear*.

**Noun idioms** are represented in our data by 65 items, for example: *a black day*; *a race against time/the clock*; *the/a sacred cow*; *the calm before the storm*; *an unwritten law*; *the other side of the coin*; *the silent majority*. Several of these nominal phrases can turn into verbal phrases or sentence idioms when a non-obligatory verb is added (mainly *to be*), cf. *(to be) one’s right hand*; *(to be) a wolf in sheep’s clothing*; *(this is) the beginning of the end*.

**Adverbial idioms.** We came across 25 widespread idioms with the function of an adverbial phrase. Examples include *at full steam*; *behind the scenes*; *from the cradle to the grave*; *in the twinkling of an eye*; *night and day*, and *under the seal of secrecy*. Again, this assignment can be changed by adding a verb: *(to be) armed up to the teeth*; *(to run/approach) with seven-league boots*.

**Adjective idioms** are generally very rare. The actual existence of idioms with an adjectival function must be questioned because they are mostly used in predicative position together with the verb *to be* and thus cannot be distinguished from verb idioms. Our collection contains ten widespread adjective idioms with the specific structure of comparisons, e.g. *as poor as a church mouse*; *as hungry as a wolf* or *as old as Methuselah*. They are all similes: the meanings of the adjectives remain literal but they are intensified by the comparative structure; this leads us to the next observation.

Two subgroups of figurative units can be singled out due to their salient structures, namely similes and (irreversible) binomials. Both groups are considered to belong to the category of idioms, and both appear in the various syntactic functions of idioms (as verbs, nouns and adverbs).

**Similes.** The structure of similes consists of the topic of comparison (*poor* in the first example above) and the vehicle of comparison (*church mouse* in the example), which is often a noun phrase. The two parts are connected via a particle (*as* or *like*). Apart from the similes that include an adjective, there are about 13 widespread similes in our data collected for English which belong to the grammatical class of verbs, e.g. *to collapse like a house of cards*; *to be dropping like flies*; *to shoot up like mushrooms*. Some of these can be used adverbially, cf. *(to come unexpectedly) like a bolt from/out of the blue*.

The multilingual approach, however, qualifies this picture, since some languages prefer the comparative structure where other languages do not use a simile. Latvian is well-known for making extensive use of the particle *kā* ‘as, like’, cf. *to be a thorn in someone’s eye* vs. Latvian *kā skabarga acī* ‘like a splinter in the eye’, *to treat someone with kid gloves* vs. Latvian *kā ar zīda cimdiem* ‘like with gloves of silk’ or *to be sitting on a powder keg* vs. Latvian *sēdēt kā uz pulvera mucas* ‘to sit like on a powder keg’.

**Binomials**, also called *irreversible binomial idioms*, *binomial pairs* or *word-pairs*, are sequences of two or more constituents that belong to the same grammatical category, have some semantic relationship and are joined by a conjunction like *and* or *or*. Our data contain 15 widespread binomials, among them those in the function of a verb (*to move heaven and earth*), a noun (*one’s own flesh and blood*), and an adverb (*from head to foot*). Binomials are often “irreversible” in one language (cf. Malkiel 1959). Cross-

linguistically, however, we may encounter a reverse order of the constituents, cf. *to be all skin and bones* vs. Finnish *olla pelkkää luuta ja nahka* “to be only bone and skin”; *(to be/fight) like cat and dog* vs. German *wie Hund und Katze (leben/sein)* “(to live/be) like dog and cat” or *night and day* vs. French *jour et nuit* “day and night”, etc. Occasionally, a binomial in one language may have equivalents without a binomial structure in other languages (cf. *to go through fire and water (to help someone)* vs. German *für jmdn. durchs Feuer gehen* “to go through the fire for someone”). This does not affect its status as being a widespread figurative unit.

To sum up, individual idioms and their variants may belong to more than one of the subcategories of idioms. This fact is particularly evident from a multilingual perspective and has no influence on the comparability of the idioms across different languages. For example, several languages use equivalents of the sentence idiom “the rats are leaving the sinking ship” while other languages prefer a verb idiom “to leave a sinking ship like rats”, which is at the same time a simile.

## 2.2 Cultural Foundation: Toward a Macro-structure of the Lexicon of Common Figurative Units

Part of this section will be devoted to the question of how to structure the “Lexicon of Common Figurative Units”. An arrangement of the widespread idioms according to their origins (i.e. to their cultural and historical foundation) seems to be the best solution. As there is a relevant connection between figurative language and culture, we will also discuss the concept of culture in figurative language and related phenomena (idiom motivation and idiom etymology).

### 2.2.1 The Problem of Grouping Figurative Units

There are several possibilities to present our widespread idioms in the form of a written lexicon. The obvious thing to do would be to list them alphabetically, according to the first nominal constituent, as is usual in idiom dictionaries. However, this would result in idioms being strung together that have nothing to do with each other. Using English as a starting point, idioms as diverse as the following would be listed under the letter A:

*to be conspicuous by one's absence; Achilles' heel; much ado about nothing; the alarm bells are ringing; the Alpha and Omega; to say amen to everything; to run amok; to welcome/greet someone with open arms; to sweep the Augean stables.*

These idioms go back to Greek mythology, the Bible, post-classical literary works, a semiotized gesture and to the domains of military and foreign culture, respectively. Questions that would remain unanswered would include



the placement of variants of constituents (*to be sb.'s right hand/arm*) and idioms that do not have an English equivalent (cf. German *im Adamskostüm* “in **Adam's** costume”). Such an alphabetic arrangement would cause needless repetitions and redundancies.

Another arrangement could start from the main idiom constituents or more precisely from the significant meanings they have outside of their idiomatic context. For example, we could group those idioms together that have a BODY PART as one of their constituents. Because EYE is a frequent constituent in idioms of many languages, let us list the widespread idioms which contain this constituent:

*the scales fall from sb.'s eyes; to be a thorn in sb.'s eye; to see the mote in one's neighbor's eye (and not the beam in one's own); an eye for an eye, a tooth for a tooth; (to be able to see sth.) with the naked eye; not to believe one's eyes; to cast one's eyes over sth.; to close/shut one's eyes to sth.; to hit sb. in the eye; to open sb.'s eyes; to throw/cast dust into sb.'s eyes; in the twinkling of an eye; (from) a bird's-eye view.*

Again, we would deal with rather diverse idioms. The origins of the first four idioms lie in the Bible and other ancient texts while the last idiom does not refer to the eye of a human being. Again, we face the problem where to place those idioms which have no English equivalent, cf. German *mit einem lachenden und einem weinenden Auge* “with one laughing and one weeping **eye**”, whose wide circulation is ascribed to Shakespeare's verses and older literary sources (but is not an idiom of modern English), and German *etwas wie seinen Augapfel hüten* “to take care of sth. like one's **eyeball**/apple of the **eye**” which is of biblical origin.

This kind of arrangement is well known in phraseological studies. It is the traditional grouping of idioms into “thematic groups”, starting from the domains to which the constituents belong (e.g. ANIMALS, PLANTS, COLORS, GARMENTS, BODY PARTS, etc.), sometimes mistakenly referred to as “onomasiologic” approach. Such classifications have often led to unstructured collections of idioms of most diverse origins, disregarding their culture-based nature as an essential feature.

Let us look at a few extreme cases, for example when the idiom “to carry **owls** to Athens” is subsumed under BIRDS and the idioms *God's lamb* and *to worship the golden calf* under FARM ANIMALS (Jarosińska 1991: 251–253), when equivalents of *to weep/shed crocodile tears*, which go back to old legends, are merely grouped under WILD ANIMALS (cf. Sabban 2007: 592, among others), or the idioms with the constituent *laurels* (e.g. *to gain one's laurels*, which originates from a popular custom in antiquity), are considered to belong to the thematic group of PLANTS because LAUREL is a plant, outside of the idiomatic context (Gréciano 2002b: 439). Along the same lines, Mejri (2004: 67) classifies the French idiom *jeter le gant* (a widespread idi-

om, cf. English *to throw down the gauntlet*) as belonging to the thematic group of GARMENT, together with COAT, SHIRT, SHOES, etc., instead of taking its cultural historic background into account (it refers to a gesture of medieval knighthood, a sign of issuing a challenge).

The same holds true for PARTS OF THE BODY. Although there has been a long tradition of analyzing idioms containing somatic constituents (e.g. Rajchštejn 1980: 91f), Braun and Krallmann (1990: 79) are surprised by the multitude of idioms referring to the human body. Their list includes idioms such as *to nourish a viper in one's bosom*; *to be a thorn in someone's eye*; *to have feet of clay* etc. where the "body parts" are just not of importance when compared to these idioms' origins in well-known, identifiable textual sources. Similarly, traditional idiom research is prone to subsume idioms that refer to HEART under BODY PARTS, together with e.g. idioms referring to FEET, NOSE, or SHOULDER, although idioms that use semiotized concepts like HEART are of completely different origin.

These examples show that it would make no sense to start from the meanings of the idiom constituents (i.e. grouping them "thematically") in order to arrange the idioms of the "Lexicon of Common Figurative Units" in a meaningful way.

Finally, another possible principle of structuring the "Lexicon" should be considered here briefly, that of starting from the images evoked by the idioms' overall lexical structure. In this case, the idiom *to (always) fall/land on one's feet* would not be subsumed under BODY PARTS (based on the element of *feet*) but under ANIMAL BEHAVIOR. The idiom is based on the observation of a cat's behavior, with the animal's remarkable ability to land on its paws after falling from a great height. About a dozen widespread idioms literally refer to the behavior of animals, cf.

*to (always) fall/land on one's feet*; *to lick one's wounds*; *to show one's teeth to sb.*; *to prick one's ears*; *with one's tail between one's legs*; *to make one's mouth water*; *to play cat and mouse with sb.*; *(to be/fight) like cat and dog*; *to howl with the wolves*; *like rats leaving/deserting the sinking ship*; *to die/to be dropping like flies*, etc.

Only some of these idioms refer to the animals by name, but observations of the animals' behavior and reflexes can be recognized in the imagery of the other idioms as well: of a dog or a cat licking an injured part of their body, of a dog or wild animal baring their teeth, of a dog or horse raising their ears at a sudden noise, of a frightened dog slinking away humiliated, and so on. This arrangement of idioms is actually practicable for some domains and we will make use of it to some extent. However, there are several overlaps and inconsistencies as well.

A number of idioms whose literal meaning refers to animals and their alleged behavior do not owe their existence and wide distribution primarily to real experiences with the animals but to long literary traditions, ranging from old semiotizations since antiquity (as is the case with *crocodile tears* mentioned above, with similes like *as hungry as a wolf*, *as gentle as a lamb*, cf. 2.1.5, and also with an “inconspicuous” idiom like *to prick one’s ears*) to idioms promoted through once well-known narratives (e.g. *(to be/fight) like cat and dog*) or once famous proverb anthologies (e.g. *with one’s tail between one’s legs* and *to howl with the wolves*).

For a reasonably coherent arrangement of the data in the “Lexicon of Common Figurative Units”, we have found another solution. We will group the widespread idioms according to their origins, as far as this is possible in accordance with the current state of research. Consequently, it is an important step to determine the “true etymology” of each individual idiom. This makes it necessary to first go into the concepts of culture, idiom motivation, and idiom etymology in more detail.

### 2.2.2 Culture: A Constant Element of Figurative Units

In earlier stages of idiom research, the interest in *cultural phenomena* was varied. Until quite recently, topics like idiom syntax, idiom semantics, idiom pragmatics, including sociolinguistic and psycholinguistic perspectives, text-related modifications of idiom structure and cross-linguistic research, have figured more prominently in many studies from Europe than cultural topics. Current studies, however, regard *culture* as a *fundamentally acknowledged constant* in phraseology. They have highlighted the fact that phrasemes (idioms in particular) are not only units of a sign system, language, but also carriers of cultures, pointing out the necessity for modern idiom research to turn to cultural phenomena (cf. e.g. Cowie 1998; Dobrovol’skij 1998: 55-58; Teliya et al. 1998; Gréciano 2002; Sabban 2007, 2008; Dobrovol’skij/Piirainen 2005; Piirainen 2007b, 2008a, 2012c).

This turn toward cultural phenomena is connected not least to cognitive approaches to idioms (and other figurative units of the lexicon). The cognitive perspective allows us to place the relevant knowledge structures which underlie idioms (conceptual structures such as frames and scripts, textual knowledge, symbolic knowledge, i.e. a wide range of cultural knowledge) in the center of idiom research.

Despite this general agreement on the fact that culture plays an important role for most questions concerning idioms, only a few studies have actually treated the relation between figurative units and culture in detail. While it is true that titles of phraseological studies rarely use the words *culture* or *cultural* (judging from relevant bibliographies), the studies themselves make

ample use of these terms (as well as of words such as *culture-specific*, *culture-based*, *culturally marked*, *culturally significant*, *culturally tainted*, *culturally bound*, *cultural connotations*, etc.). Due to the vagueness of the term *culture-specificity*, Sabban (2007) proposes replacing it with the term *culture boundness*. It should be added that the term *cultural foundation* has been applied when investigating cultural phenomena that underlie conventional figurative units (cf. Dobrovol'skij/Piirainen 2005: 216–251).

There are numerous definitions of *culture* in linguistics, anthropology and related cultural studies. As early as in 1952, Kroeber and Kluckhohn list no less than 164 definitions of culture from popular and academic sources, and we certainly do not want to add a new one to them. Notions of culture seem to waver between a wide and a narrower concept. Wierzbicka (1992, 1996) for example, favors a wide concept, pointing out that almost everything in language reveals a certain degree of cultural specifics. She states that the meanings of most words differ from language to language because they are cultural artifacts, reflecting aspects of the cultures in which they were created (1996: 15f).

Other principal characteristics of culture come to the fore in the field of semiotics. Central to the attempts at defining culture is here the human predisposition to create signs and to give significance to all things that surround us; culture is viewed as a system of symbols or meaningful signs. Some ideas of this semiotic view of culture have been adopted by Dobrovol'skij and Piirainen (2005) to describe figurative language. The authors' concept of culture focuses on the cognitive nature of culture and on collective conceptions as they are fixed in language so that they advocate the notion of culture as "the sum of all ideas about the world (including fictional, mythological, etc. ideas) that are characteristic of a given community" (2005: 213).

### 2.2.3 Motivation of Figurative Units

All of the 380 widespread idioms we have identified are clearly motivated. This means that most speakers of the various languages of our project can activate certain knowledge structures to make sense of the use of a given idiom in the meaning conventionally ascribed to it. The relationship between the two conceptual levels, between the mental image evoked by the lexical structure and the figurative meaning, becomes comprehensible to them. This semantic transparency is certainly an important prerequisite for the wide dissemination of these idioms. Opaque figurative units, where the link between the two conceptual levels is not clear (as is the case with the idiom *spick and span*, figuratively meaning 'very clean, neat, tidy') would have no chance to spread into many other languages.

Since motivation is of great importance for analyzing and describing the cultural foundation of figurative units, let us look at this issue in more detail. There are several types of motivation, i.e. several different ways how the connection between a lexical structure and its figurative meaning is understood. In general, however, most idioms are motivated semantically. Cognitive-semantic linguistics assumes that speakers processing an idiom map the *source concept*, i.e. the image evoked by the idiom's literal meaning, onto the *target concept*, i.e. the concept behind the idiom's lexicalized meaning. In European idiom research, it is common to distinguish between two principal types of semantic motivation: the *metaphoric motivation* type and the *symbol-based motivation* type.<sup>12</sup> The latter manifests itself only in a small number of widespread idioms (from the domains of color symbolism, number symbolism, and animal symbolism).

We must distinguish between two types of metaphoric motivation, which require different theoretical frameworks. The motivational links that are relevant to metaphorically motivated idioms can be explained either on the abstract, *superordinate level* of categorization or on the more concrete *basic level* (following Rosch's (1978) categorization of semantic prototypes). The description of metaphors on a very abstract level is well-known through terms like *conceptual metaphor* or *metaphoric model*, developed within the framework of the Conceptual Metaphor Theory (e.g. Lakoff/Johnson 1980; Lakoff 1987). The description of metaphors on the basic level starts from the rich imagery that is evoked by the literal reading. Here, detailed knowledge about the source frames is involved in the interpretation. This frame does not have to be fixed as a whole in the lexical structure of the idiom but may be evoked by a relevant slot. We call these kinds of metaphors *frame-based metaphors*.<sup>13</sup>

There are also a few minor motivation types, e.g. so-called *kinetic idioms*, *play on words* and the *indexal motivation* (cf. references in footnote 12), all of which are not important for our widespread idioms research. In contrast to that, another, not purely semantic type of motivation, namely *intertextuality*, is of outmost importance for our study. Intertextuality is understood here as the relation of conventional figurative units and existing texts (see 2.3 below). The matter is complicated by the fact that intertextuality can be regarded as a type of motivation only for some individual figurative units. In most cases, the idioms' originating from pre-existing texts does not contribute to their motivation but is merely a question of their etymology. The domain of

<sup>12</sup> See Dobrovol'skij/Piirainen (2005: 87–101, 2009: 17–41, 2010); Burger 2007; Piirainen 2011a, among others. For linguistic motivation in general see e.g. Radden/Panther 2004.

<sup>13</sup> See Baranov/Dobrovol'skij (1996, 2008); Dobrovol'skij/Piirainen (2005: 161–185) for more details.