

Arabic Idioms

A corpus-based study

Ashraf Abdou



Routledge Arabic Linguistic Series

Arabic Idioms

Idioms represent a fascinating linguistic phenomenon that has captured the attention of many linguists for decades. The ubiquity of these expressions in language use, the wide range of functions they perform in discourse, the problems they often cause in domains such as foreign language learning and translation, and their typical divergence from the normal rules of grammar and semantic compositionality are among the main reasons for this scholarly interest.

This book is a corpus-based study of idioms in Modern Standard Arabic. Examining Arabic idioms with regard to their semantic, discursive, lexical, and grammatical properties, the author sheds light on their intricate nature, establishes the major patterns of their linguistic behavior, and provides explanations for these patterns.

Adopting a descriptive framework and systemically accounting for major linguistic phenomena, this analysis will be accessible to linguists, translators, lexicographers, translation software developers, and language teachers.

Ashraf Abdou is Adjunct Assistant Professor of Teaching Arabic as a Foreign Language at the American University in Cairo and Lecturer of Linguistics at Cairo University. His research focuses on Arabic phraseology, Arabic corpus linguistics, spoken discourse analysis, Arabic lexicography, and English–Arabic contrastive phraseology.

Routledge Arabic Linguistics Series

General Editor: Clive Holes (University of Oxford)

Editors:

El-Sa'id Badawi (American University in Cairo)

Adrian Gully (University of Melbourne)

Yasir Suleiman (University of Cambridge)

Keith Walters (University of Texas)

James Dickins (University of Leeds)

Jonathan Owens (Bayreuth University)

The *Routledge Arabic Linguistics Series* publishes high quality, academically rigorous research on Arabic linguistics to two main readerships: non-Arabist general linguists with an interest in Arabic, and students and researchers already in the field of Arabic language and linguistics. Both synchronic and diachronic studies of Arabic are welcome which aid our understanding of the historical evolution and the present state of Arabic, whether dialectal or standard. Works written from a sociolinguistic (e.g. language variation), socio-historical (e.g. language history), sociological (e.g. language planning), or psycholinguistic (e.g. language acquisition) perspective are welcome, as are studies of Arabic stylistics, pragmatics, and discourse analysis. Descriptive dialectological works also fall within the scope of the Series, as do works which focus on the evolution of mediaeval Arabic linguistic thought. Proposals or scripts for the Series will be welcomed by the General Editor.

PREVIOUSLY PUBLISHED BY CURZON

1. Medieval Islamic Pragmatics

MSunni legal theorists' model of textual communication

Mohamed Mohamed Yunis Ali

2. Jewish and Muslim Dialects of Moroccan Arabic

Jeffrey Heath

3. Language Contact and Language Conflict in Arabic

Variations on a sociolinguistic theme

Edited by Aleya Rouchdy

PUBLISHED BY ROUTLEDGE

4. Structure and Function of the Arabic Verb

Maher Bahloul

5. Arabic in the City

Issues in dialect contact and language variation

Edited by Catherine Miller with Enam Al-Wer, Dominique Caubet and Janet Watson

6. Information Structure in Spoken Arabic

Edited by Jonathan Owens and Alaa Elgibali

7. Islamist Rhetoric

Language and culture in contemporary Egypt

Jacob Høigilt

8. Arabic Idioms

A corpus-based study

Ashraf Abdou

Arabic Idioms

A corpus-based study

Ashraf Abdou

First published 2012
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
710 Third Ave, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2012 Ashraf Abdou

The right of Ashraf Abdou to be identified as author of this work has been asserted by him in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Abdou, Ashraf.

Arabic idioms: a corpus based study / Ashraf Abdou.

p. cm.—(Routledge Arabic linguistics series)

Includes bibliographical references and index.

1. Arabic language—Idioms. 2. Arabic language—Spoken Arabic. 3. Arabic language—Discourse analysis. I. Title.

PJ6167.A23 2011

492.7'5—dc22

2011003469

ISBN: 978-0-415-60340-9 (hbk)

ISBN: 978-0-203-80878-8 (ebk)

Typeset in Times New Roman
by RefineCatch Limited, Bungay, Suffolk

To my family, for their love and continuous support

To Sara, with great hopes for the future

Contents

<i>Foreword</i>	xi
<i>Acknowledgments</i>	xiii
<i>List of abbreviations</i>	xiv
<i>List of tables</i>	xvi
<i>Transliteration scheme</i>	xvii
1 Introduction	1
1.1 <i>Significance of the study</i>	1
1.2 <i>Research questions</i>	2
1.3 <i>Modern Standard Arabic</i>	3
1.4 <i>Sources of idioms in MSA</i>	3
1.5 <i>Arabic idioms and other types of multiword units</i>	4
1.6 <i>Using corpora in idiom studies: advantages and limitations</i>	7
1.7 <i>Model of analysis</i>	9
1.8 <i>Structure of the book</i>	11
1.9 <i>Conventions used in this book</i>	11
2 Methodology and data	13
2.1 <i>Collecting Arabic idioms</i>	13
2.2 <i>Classifying Arabic idioms</i>	16
2.3 <i>Selecting the sample</i>	17
2.4 <i>Corpus data: sources, collection procedures, and preparation</i>	18
2.5 <i>The language of Arabic newspapers</i>	24
2.6 <i>Data examples and their frequency</i>	26
3 The semantic structure of Arabic idioms	37
3.1 <i>Motivation</i>	37
3.2 <i>Patterns of semantic extension in idioms</i>	39
3.3 <i>Arabic idioms and isomorphism</i>	55
3.4 <i>Discussion</i>	61

x	<i>Contents</i>	
4	The discursive behavior of Arabic idioms	66
	4.1 <i>Discursive functions of idioms</i>	66
	4.2 <i>Idioms and cohesion</i>	91
5	The lexical and grammatical behavior of Arabic idioms	102
	5.1 <i>Systematic variation vs. wordplay</i>	102
	5.2 <i>Lexical variation</i>	104
	5.3 <i>Lexicogrammatical behavior of idioms</i>	109
	5.4 <i>The grammatical behavior of idioms</i>	113
6	Conclusions	150
	6.1 <i>Major findings and their interpretations</i>	150
	6.2 <i>Contexts of use and applications</i>	155
	6.3 <i>Limitations of the study</i>	156
	6.4 <i>Ideas for future research</i>	156
	<i>Notes</i>	158
	<i>Bibliography</i>	160
	<i>Index</i>	165

Foreword

Idioms have attracted increasing attention from linguists over the past few decades, from semantic, discursal, cultural and other perspectives. With advances across the language sciences, our understanding of what idioms are and what they do has become much richer and more subtle. Corpus linguistics has also developed exponentially, as computers have become able to handle ever huger amounts of data, and the computational tools available to linguists have grown in sophistication.

There have been numerous studies of idioms in English, including a few which have been corpus-based, such as Rosamund Moon's 1998 book *Fixed Expressions and Idioms in English: A corpus-based approach*. There are now a good number of corpora for Arabic, covering both the standard language and colloquial dialects. Studies of idioms in Arabic are, however, rare, the only work previously published in the West being L.I. Torlakova's 1998 paper "Some cultural and ethnic elements in Modern Standard Arabic idioms".

Ashraf Abdou's book *Arabic Idioms: A corpus-based study* is groundbreaking, in that it adopts a sophisticated multi-faceted theoretical approach to idiom analysis, combined with a painstaking use of one of the largest publicly available Arabic corpora, *Arabic corpus* (over 83 million words). The book looks at Arabic idioms from a number of perspectives: syntactic structure (Chapter 2), figurative semantic structure, including the very interesting issue of the way in which metaphor and metonymy sometimes interact with one another in idioms (Chapter 3), discursive behavior (Chapter 4), cohesion (Chapter 4), and lexical and grammatical behavior of idioms (Chapter 5). Because of its corpus orientation, *Arabic Idioms: A corpus-based study* is able to say not only what exists, but what typically exists – what are the standard (as well the exceptional) features of Arabic idioms.

This book will, of course, be of particular interest to Arabists. However, as the first study to provide both broad and profound insights into idioms in a non-Western language, it is of great significance to the theoretical study of idioms more generally. The book also has practical implications. The teaching of Arabic lexis to English-native learners (and no doubt learners of other native languages) has traditionally concentrated almost exclusively on single-word units; this book reveals how important idioms, and other multi-word units, are in Arabic, demonstrating

the need to focus more on these in Arabic teaching. Similarly, this work shows that the scant attention traditionally paid by Arabic/English dictionaries to idioms (and other multi-word units) will need to give way – particularly as Arabic/English lexicography becomes progressively based on Arabic (as well as English) corpora – to an approach which fully recognizes the importance of these phenomena.

Professor James Dickins
Head of Arabic and Middle Eastern Studies
University of Leeds

Acknowledgments

I would like to thank the series editor, Professor Clive Holes, for his support and encouragement throughout the production of this book. I would also like to thank the editing team, Joe Whiting and Suzanne Richardson, for their consonant support and endless patience.

As the origin of this work can be traced to my doctoral dissertation, I would like to thank Dr Paul Bennett, Professor John Payne, Professor Harold Somers, and Dr Martina Faller at the University of Manchester and Professor James Dickins at the University of Leeds for their valuable guidance and advice. I am thankful to Professor Dilworth Parkinson at Brigham Young University for the help he provided me and for making available the main corpus I draw on in this work.

I am deeply indebted to Professor El-Said Badawi at the American University in Cairo. No words can express my gratitude for his guidance, support, and encouragement.

I am grateful to my family for their support and patience. Without this I simply would not have made it to this point. I also would like to thank my friends, especially Mahmoud Al-Herthani at Al-Aqsa University in Gaza, Ahmed Al-Emam at the University of St Andrews, and Latifa Shamsan at the University of Manchester for the useful discussions I had with them during my stay in Manchester.

Finally, I am grateful to Continuum International Publishing Group for granting me the permission to use major parts of the paper I wrote for their book *Perspectives on Formulaic Language: Acquisition and communication* in the third chapter of this book.

Abbreviations

1	first person
2	second person
3	third person
ACC	accusative
ADJ	adjective
ADV	adverb(ial)
APTCP	active participle
CAUS	causative
COMP	complementizer
COND	conditional
DEF	definite
DEM	demonstrative
DIMIN	diminutive
DU	dual
EMPH	emphasizer
F	feminine
FUT	future
GEN	genitive
HUM	human
IMP	imperative
IND	indicative
INDF	indefinite
JUSS	jussive
M	masculine
MOD	modifier
NEG	negation, negative
NOM	nominative
NONHUM	nonhuman
OBJ	object
PASS	passive
PL	plural
POSS	possessive
PRS	present

PST	past
PPTCP	passive participle
Q	question particle
REFL	reflexive
REL	relative
SBJ	subject
SBJV	subjunctive
SG	singular
VOC	vocative

Tables

2.1	Syntactic classes of Arabic idioms	16
2.2	The word count of each newspaper in <i>Arabicorpus</i>	18
2.3	Components of <i>Arabicorpus</i> other than newspapers	19
2.4	Verb-subject idioms	27
2.5	Verbal idioms	28
2.6	Nominal idioms	31
2.7	Prepositional idioms	32
2.8	Adjectival idioms	33
2.9	Adverbial idioms	34
3.1	Distribution of idioms over figurative patterns	40
3.2	Isomorphism and the syntactic types of idioms	57
3.3	Isomorphism and the patterns of semantic extension in idioms	57
4.1	Distribution of idioms over discursive functions	69
5.1	The syntactic structures of verb-subject idioms	128
5.2	The syntactic structures of verbal idioms	129
5.3	The syntactic structures of nominal idioms	131
5.4	The syntactic structures of prepositional idioms	132
5.5	The syntactic structures of adjectival idioms	132
5.6	The syntactic structures of adverbial idioms	132
5.7	Distribution of the relevant idioms over the types of embedding	141

Transliteration scheme

Consonants

Arabic	Transliteration	Arabic	Transliteration
ء	'	ض	d
ب	b	ط	t
ت	t	ظ	z
ث	th	ع	'
ج	j	غ	gh
ح	h	ف	f
خ	kh	ق	q
د	d	ك	k
ذ	dh	ل	l
ر	r	م	m
ز	z	ن	n
س	s	ه	h
ش	sh	و	w
ص	ṣ	ي	y

Short and long vowels

Arabic	Transliteration	Arabic	Transliteration
ا	a	أ or إ	ā
ي	i	و	ū
و	u	ي	ī

Notes on transliterations

- For the most part, this scheme follows the style used by *The International Journal of Middle East Studies*.
- The definite article *al* is represented as is without showing any sort of assimilation.
- Phonological junctural features in word-final position are represented in the transliterations preceded by a hyphen, but they are not represented in the glosses.

- When appropriate, the subscripts $_i$ and $_j$ have been placed adjacent to the nouns and/or pronouns in the examples to indicate co-referentiality.
- In the transliterations, sometimes irrelevant parts in the corpus examples have been omitted and replaced with [. . .].
- I have not used the transliteration system adopted here in giving the titles of newspapers, websites, etc. I have mainly used the titles as they are given by the newspaper or website, or as written by the developer of the corpus I draw on.
- When appropriate, the two Arabic words *shakhs* “person” and *shay*’ “thing” are used in the transliterations to stand for any elements that are supplied by the co-text to fill in open slots in the structure of the idiom. These elements are represented in the glosses, literal translations, and proposed English translations by the abbreviations *sb* and *sth* for “somebody” and “something”, respectively. Italics have been used in all of these cases.

1 Introduction

Idioms represent a fascinating linguistic phenomenon that has captured the attention of many linguists for decades. The ubiquity of these expressions in language use, the wide range of functions they perform in discourse, the problems they often cause in domains such as foreign language learning and translation, and, very importantly, their typical divergence from the normal rules of grammar and semantic compositionality are among the main reasons for this scholarly interest.

This book is a corpus-based study of idioms in Modern Standard Arabic. Examining the data with regard to their semantic, discursive, lexical and grammatical properties, the study sheds light on their intricate nature, establishes the major patterns of their linguistic behavior, and provides explanations for these patterns.

1.1 Significance of the study

A corpus-based study of Arabic idioms may have considerable significance both on the theoretical and practical levels. As far as the former is concerned, the study can help gain an appreciation of the position of this phenomenon in the language's structure and how it interacts with other lexical and grammatical aspects.

In the last few decades, phraseology, the domain of linguistic enquiry interested in studying prefabricated constructions, has acquired a central position in the scientific examination of language (see Wray 2008; Langlotz 2005). Granger (2005) points out that the ubiquity and centrality of phraseology is supported by the findings of corpus-based research and psycholinguistic studies “which present holistic storage as the default type of processing” (p. 166). Wray (2002) emphasizes that “[n]o model of language which includes a notional lexicon can avoid storing in it morphemes on the one hand and irregular words and word strings on the other” (p. 264). Also, Cowie (1998b: 2–3) notes that:

[t]he notion that native-like proficiency in a language depends crucially on a stock of prefabricated units – or “prefabs” – varying in complexity and internal stability is now set in critical opposition to the atomistic view, rooted in generative theory, that the workings of a language can be explained by a system of rules of general applicability, a lexicon largely made up of minimal units and a set of basic principles of semantic interpretation.

2 Introduction

Furthermore, with regard to idioms in particular, because of the central role that different patterns of semantic extension play in their creation, a detailed corpus-based study of this type of multiword units may offer useful insights in the context of metaphor studies.

As to the practical aspect, due to their semantic non-compositionality and, sometimes, grammatical irregularity, idioms are infamous for the thorny problems they cause not least in translation, natural language processing, and language learning. A detailed study of the corpus occurrences of idioms seems essential for developing more effective ways to handle such problems. For instance, it is expected that the linguistic descriptions provided throughout the present work can be encoded, using appropriate computational-linguistic formalisms, within different natural language processing applications that involve Arabic.

Studies of idioms can also provide insights into the cultural mentality of the speech community. Teliya *et al.* (1998) point out that phraseology amply demonstrates the strong relationship between language and culture. To them, it can be regarded as a storehouse of cultural data, that is, a repository of information on the prominent values, attitudes, and ideas in the speech community. They also note that such “[c]ultural connotations are especially vivid in idioms and restricted lexical collocations” (p. 59).

Even though this work does not adopt a linguo-cultural approach, it still can yield some insights into the Arab cultural mentality as it manifests itself in the meanings of Arabic idioms and, particularly, in the evaluative content they often communicate (for an analysis of the cultural elements in some Arabic idioms, see e.g. Torlakova 1998). Gathering cultural information that relates to the form and meaning of different linguistic units is important in the context of practical applications such as lexicography, language learning, and translation (see Teliya *et al.* 1998).

In the context of a corpus-based study of German idioms and collocations, Fellbaum (2007b) refers to both the theoretical and lexicographic aspects of the issue, emphasizing that these expressions “are *frequent, non-marginal* phenomena that pose a *challenge to our understanding of grammar and lexis*; the lack of rich data and adequate linguistic theories is reflected in insufficient *lexicographic treatment*” (p. 2, her italics).

1.2 Research questions

This book attempts to provide answers to various questions regarding the linguistic behavior of Arabic idioms. The most prominent points that are under examination are listed below according to the area of linguistic enquiry they belong to:

- As to the semantic properties of idioms, two major points are examined: what are the patterns of semantic extension that underlie the production of Arabic idioms? And, how can these idioms be classified with respect to the notion of isomorphism?

- Regarding the discursive behavior of Arabic idioms, the study investigates both their discursive functions and the ways in which they contribute to the cohesion of their texts.
- With reference to the lexical and grammatical behavior of the data, this work is interested mainly in their lexical variation, changes in the lexicogrammatical complexity, perspective-adaptation, inflectability, active and passive voice, syntactic types and structures, constituent order, embedding, and predicatization.

1.3 Modern Standard Arabic

In characterizing Modern Standard Arabic (MSA), Holes (2004: 5) writes that it:

is the modern descendant of Classical Arabic [for short CLA, i.e. mainly the language of the Islamic scriptures and classical Arabic literature], unchanged in the essentials of its syntax but very much changed, and still changing, in its vocabulary and phraseology. This unified, codified pan-Arab variety of Arabic is used for virtually all writing in the Arab world and nowadays, in its spoken form, also dominates the airwaves and the television channels of every Arab country. As the normal medium for formal discourse, it is used in all news broadcasts, political speeches, official announcements, and – most crucially – education in every Arab country.

MSA can be used productively, with varying degrees of grammatical accuracy, by literate native Arabic speakers (Badawi 1995). As Ryding (2005: 7) points out, this variety:

serves not only as the vehicle for the current forms of literature, but also as a resource language for communication between literate Arabs from geographically distant parts of the Arab world. A sound knowledge of MSA is a mark of prestige, education, and social standing; the learning of MSA by children helps eliminate dialect differences and initiates Arab children into their literary heritage and historical tradition. It aids in articulating the connections between Arab countries and creating a shared present as well as a shared past. Education in the Arab countries universally reinforces the teaching and maintenance of MSA as the single, coherent standard written language.

Last but not least, MSA plays a major role in intercultural communication, inasmuch as most of the translation into and from Arabic takes place using this variety.

1.4 Sources of idioms in MSA

Idioms in MSA come from several sources. For example, some of them clearly originate in CLA and can be found, sometimes with different meanings, in classical Arabic dictionaries. Second, many contemporary Arabic idioms have

4 Introduction

occurred as a result of the intensive translation process that has been going on for decades from Western languages, in particular English, into Arabic, especially in the (print) media (see e.g. Stetkevych 1970). This process has expanded the phraseological repertoire of MSA with many examples of loan translations. As Holes (2004: 315) notes:

[m]uch of the news reporting in the Arab media is in the form of rapidly produced and often rather literal translations of English or French language news agency reports. In this way, quantities of new phrases are coined ad hoc by journalists and thence find their way into everyday use without ever having received the endorsement of the [Arabic language] academies.

Finally, spoken Arabic dialects also serve as a source of idioms for MSA. However, in this case, the idiom may need to undergo some lexical, grammatical, and/or phonological modifications in order for it to adhere to and become fully integrated in the MSA system. Indeed, many idioms from these sources are among the examples examined in the following chapters.

El-Said Badawi (personal communication) notes that MSA relies heavily on such sources in developing its repertoire of idioms. Two facts may constitute together a plausible explanation for this phenomenon. First, idioms often start life as creative, figurative uses of language and they are typically connected with spontaneous ways of expression (see Carter 2004 and Chapters 2 and 3 in this book). Second, MSA is confined in its spoken form chiefly to formal situations, and, on the whole, is used creatively by only highly educated Arab writers and poets.

However, that is not to say that the users of MSA never employ its lexical and grammatical resources to produce new idioms. For instance, originally one-off figurative expressions that occur in contemporary Arabic literature and which cannot be traced to any of the sources outlined above may sometimes find their way into everyday language and become part of the phraseological stock of MSA through repeated use.

1.5 Arabic idioms and other types of multiword units

For the purposes of this study, the term *idiom* is understood as: a multiword unit that has a syntactic function within the clause and has a figurative meaning in terms of the whole or a unitary meaning that cannot be derived from the meanings of its individual components.

This characterization draws heavily on some classifications of multiword units in the literature of phraseology, e.g. Howarth (1996). However, it subsumes phenomena that do not appear to be part of his understanding of what idioms are. A case in point is syntactically ill-formed expressions. To Howarth, idioms belong to a category of multiword units that he terms *composite units*, which essentially have a syntactic function in the clause. In characterizing composite units, he notes that “they are exponents of recognizable syntactic units (grammatically

well-formed)” (p. 37). In contrast, any Arabic expression that meets the criteria stated in the foregoing characterization but shows some grammatical irregularity has been considered an idiom in the present study.

For an expression to be considered a multiword unit, it should exhibit two essential qualities: institutionalization and relative fixedness. Institutionalization is a sociolinguistic process during which the expression becomes more current in use (Barkema 1996). Institutionalization therefore can be established by the frequency of the expression. However, it is appreciated that many multiword units are not frequent in some corpora due to their restricted domains of use (Moon 1998). Therefore, “we must not rule out the possibility that an utterance which does not occur repeatedly is a formula” (Hickey 1993: 33, and for a discussion of the relationship between frequency and formulaicity, see e.g. Wray and Perkins 2000 and Howarth 1998).

Relative fixedness refers to the fact that the expression does not allow for certain lexical and/or grammatical changes while preserving its conventionalized meaning. This could involve restrictions, for instance, on its constituent order, voice, inflectability, and/or lexis substitutability. Such restrictions cannot be explained by the general grammatical rules of the particular language. In some cases, however, the restrictions may be intentionally violated so as to produce ambiguous constructions and/or to create humor (see Langlotz 2006 and Moon 1998).

The foregoing characterization of *idiom* excludes from the scope of this study several types of multiword units. The most prominent of these are *restricted collocations* and multiword units that occur outside the clause or as independent utterances, e.g. discourse structuring devices and proverbs.

Restricted collocations are of different types. For instance, they include combinations in which one component is used literally and the other is semantically depleted, as in support-verb constructions. They also include combinations in which one component is used literally and the other is used in a figurative meaning (Howarth 1996). The English examples *make a decision* and *white lie* represent these types, respectively. Some corresponding Arabic examples are:

- | | | |
|-----|--|--------------------------------|
| (1) | ittakhadha
take.M.3SG.PST
he took a decision | qarār-a-n
decision-ACC-INDF |
| (2) | kadhibat-u-n
lie.F-NOM-INDF
white lie | bayḍāʾ-u
white.F-NOM |
| (3) | ghasīl-u
washing-NOM
money laundering | l-ʾamwāl-i
DEF-money.PL-GEN |

The other major category of Arabic multiword units that is excluded from the scope of this work is expressions that occur outside the clause or as independent utterances. These are sometimes termed *functional expressions* (Cowie 1988),

(for similar approaches see Strässler 1982 and Moon 1998). Finally, linguists disagree on whether such sets belong to the field of phraseology (see e.g. Granger 2005).

In the light of the discussion above, only examples such as the following are considered as idioms in this work:

- (11) taḥta l-mijhar-i
 under DEF-microscope-GEN
 (e.g. of a situation) being examined carefully
- (12) 'anzala l-sitār-a 'alā shay'
 lower.PST.M.3SG DEF-curtain-ACC on sth
 he brought the curtain down on sth
- (13) qalaba zahr-a l-mijann-i li-shakḥs/shay'
 turn.around.PST.M.3SG back-ACC DEF-shield-GEN to-sb/sth
 he became hostile to sb/sth after he was a friend or supporter of them/it

Given the account provided above of the contexts in which MSA is used in its spoken form, this variety does not usually serve the pragmatic roles associated with expressions such as greetings and formulas of social interaction in everyday spoken discourse. Therefore, in the context of this study, a categorization of multiword units that makes a basic distinction between those which have a syntactic function within the clause and those which have mainly a pragmatic function and occur outside it or as independent utterances is preferable to one which starts with e.g. a distinction between semantically compositional and non-compositional expressions (see Howarth 1996).

Examining idioms rather than functional expressions is more likely to reveal interesting phenomena particularly with respect to their formal variation. This assumption can be supported by the fact that the presence of the idiom within the clause entails that it interacts both semantically and grammatically with other parts in the sentence. This may be the trigger for types of formal behavior that may not be found in functional expressions that do not have a syntactic role in the clause.

Furthermore, restricting the phenomena considered as idioms to fully non-compositional multiword units that occur within the clause allows the present study to concentrate on a single, relatively homogenous class of multiword units. This should facilitate discerning any patterns in the data and help in drawing up any general rules for the linguistic behavior of these expressions.

1.6 Using corpora in idiom studies: advantages and limitations

The term *corpus* is understood here as a collection of written and/or spoken material that is stored in machine-readable form and can be used for the purposes of linguistic research.

Thompson and Hunston (2001: 15) describe how large corpora represent the discourse of a speech community and point out their advantages over intuition:

8 Introduction

In recent years, many aspects of language use that in former times were accessible only through intuition have become available for study using large, machine-stored corpora. The intuition of a language user regarding a particular lexical item is the product of tens, hundreds, or thousands of experiences of that item, scattered across years of heterogeneous language experience. In representing the discourse of a community, a very large corpus can mimic, though not of course replicate, that experience.

Indeed, using corpora can help avoid the critique sometimes directed to introspection-based analyses, that is, they are “fragmentary and involve only part of what is observable in language” (Stathi 2007: 85).

Corpus research can also generate new research questions, since the availability of concordanced data on a particular phenomenon may draw the researcher’s attention to issues that they have not thought of before (Conrad and Biber 2001).

Unlike other research methods such as introspection, corpus-based research is more suitable for obtaining quantitative information on language use. As Fellbaum (2007b) notes, “corpus investigations and quantitative methods have made possible the statement of ‘soft’ rules in terms of probabilities and replaced the hard rules that previously either ‘ruled in’ or ‘ruled out’ specific structures” (p. 5).

In the context of phraseology, corpus investigations have contributed much to our understanding of the fact that speakers tend to employ the same lexical combinations time and again (Schmitt *et al.* 2004). Moreover, corpora play the major part in statistical or frequency-based research on phraseology which adopts a corpus-driven approach to discover recurring word combinations that may or may not fit pre-established linguistic classes (Granger 2005).

Exploiting corpora in idiom studies can provide several advantages that other possible approaches may be lacking. As far as the semantic study of idioms is concerned, examining corpora may help reveal semantic properties of these expressions that could be otherwise difficult to uncover. For instance, drawing on sizeable, appropriately designed corpora increases the chances of detecting all the meanings an idiom might have.

As for the study of the discursive behavior of idioms, corpus data is a more reliable tool in examining the discursive functions of idioms and in analyzing how these functions are affected by different contextual factors. For example, in her study of the evaluative function of single and multiword units, Channell (2001: 39) points out that her work:

is intended to show how analysis of the evaluative function of a word or expression can be derived from concordanced examples extracted from a corpus of language data. I want to argue (and demonstrate) that analysis of evaluation can be removed from the chancy and unreliable business of linguistic intuitions and based in systematic observation of naturally occurring data. In doing this I will argue not only that a corpus-based analysis produces a sound description of this aspect of language, but also that it allows

observations which go beyond what intuitions can achieve, in revealing evaluative functions which intuitions fail to pick up.

Moreover, examining some other aspects of the discursive behavior of idioms, e.g. their contribution to the cohesion of their texts, does not seem feasible without using corpus data.

With respect to the study of the lexical and grammatical properties of idioms, exploiting large corpora has led to some radical changes in the way many linguists had been thinking idioms work. In particular, major changes have taken place regarding the view of idioms as semantically unmotivated chunks of language that permit no or very little lexicogrammatical flexibility. For example, Fellbaum (2007b), in her work on German idioms, stresses that the fact that corpus-based studies of idioms have yielded many examples of syntactic phenomena, e.g. topicalization and clefting, that had been ruled out by linguists “underscores the need for an empirical re-evaluation of claims concerning the uses of idioms” (p. 12).

Corpora systematically provide examples of idiom variation within context. Pulman (1993) points out that the constructions that result from applying certain formal operations to idioms may appear odd if they are considered uncontextualized.

In the case of Arabic idioms, relying only on introspection and existing dictionaries has some serious shortcomings. The former is by definition highly subjective and, moreover, can be influenced by memory. As for existing Arabic dictionaries, they are not updated regularly, and, in general, are not based on comprehensive examinations of large corpora. This gives rise to the subjectivity problem again, and it may also lead to the absence of some (new) idiomatic meanings from their entries. Due to their very nature, corpus data can help tackle these problems in a more satisfactory, objective manner. For example, as to the subjectivity problem, corpus data may be regarded as a poll of how different speakers use and perceive idioms (Riehemann 2001).

On the other hand, no corpus can include every possibility in the language. It is likely that relevant phenomena may be absent even in appropriately assembled, large corpora. As a consequence, in reporting on corpus-based studies, there is often a need for some qualifications that associate the findings exclusively with the corpus or corpora used, in order to avoid unwanted overgeneralization (see e.g. Sinclair 2005).

Despite a growing need for corpus-based research on Arabic phraseology to meet both practical and theoretical ends, and even with the availability of several suitable Arabic corpora, there is still a scarcity of this type of research. The present work takes a step to fill in this gap with respect to Arabic idioms.

1.7 Model of analysis

Tognini-Bonelli (2001) distinguishes between two types of corpus investigation: corpus-based and corpus-driven. Deignan (2005: 89) describes this distinction as follows:

[C]orpus-based research starts with existing paradigms and investigates these using the corpus. Corpus-driven research, in contrast, starts with a clean slate, with no assumptions about what will be found: it places the corpus at the centre of the process, and allows new categories and rules to emerge from study.

However, she also comments that “if a range of the work that claims to draw on corpora is examined, the clarity of the distinction starts to slip, though it is still a useful one as a guide” (p. 89).

Despite its title, this study has features of both corpus-based and corpus-driven research. It is corpus-based since it begins with selecting a type of multiword units that has been characterized in the literature rather than starting with an exploration of the patterns of word combinations in the corpus used. Yet, it is also corpus-driven because it does not seek to uphold any ideas in the literature if the data suggest that other patterns or explanations do exist. The study aims to explore how its examples are used in the corpus texts and seeks to employ their occurrences as the basis for discerning any relevant patterns and suggesting any possible accounts.

With respect to such a combination, Cowie (2006) points out that researchers in the field of phraseology appreciate the advantages of an approach that exploits naturally occurring data that are obtained from large corpora and that recognizes in its analysis the syntactic and discursive functions which multiword units carry out. Several benefits of the use of corpora have been set out above. As will be seen in the coming chapters, an advantage of the recognition of a classification of multiword units that is based on their grammatical, semantic, and pragmatic properties is the explanatory potential it provides the analyst with when trying to account for some of the phenomena in the corpus data.

On the whole, this work adopts a general, descriptive framework. This helps make its analyses accessible to many reader groups, e.g. linguists, translators, lexicographers, translation software developers, and language teachers and learners.

This said, on several occasions, I have made use of the findings of other works that belong to particular theoretical frameworks. A case in point is the insights provided by some cognitive-linguistic analyses of idioms. The study of idioms and figurative language in general has been a major interest to several works that adopt the assumptions of this framework. General works on figurative language such as Kövecses (2002) and more focused ones on idioms such as Goossens (2003), Geeraerts (2003), and, especially, Langlotz (2006) in his cognitive-linguistic model of idiom variation, have been of particular use to me.

In addition, some of the ideas in the analysis of the formal properties of idioms draw on notions such as argument structure and the effects that the lexical item can have on the grammatical characteristics of the environments in which it occurs, e.g. the different patterns of complementation many verbs have. Such notions are at the heart of the analyses carried out within several contemporary grammatical frameworks.

1.8 Structure of the book

Chapter 2 sets forth the methodology of the study. Chapters 3, 4, and 5 present and discuss the results of the corpus analysis. Finally, the main conclusions of the study are provided.

Chapter 2 discusses the sources of the data, the reasons for their selection, the data collection and selection procedures, and the role of newspaper language in the Arab countries today. This chapter also provides a list of the idioms investigated in the study along with information on their frequency in the main corpus I draw on.

Chapter 3 is concerned with the semantic structure of idioms. It concentrates on the patterns of semantic extension involved in their production and their classification with regard to the notion of isomorphism. Chapter 4 focuses on the discursive behavior of idioms. It includes two major sections. The first examines the discursive functions of Arabic idioms and the other looks at how these idioms contribute to the cohesion of their texts. Chapter 5 addresses the formal behavior of the data. In three main sections, it investigates their lexical, lexicogrammatical, and grammatical behavior.

The order of the analytical chapters has been motivated by a wish to determine the extent, and facilitate the discussion, of any relationships between the semantic and discursive properties of idioms and their lexical and grammatical behavior (for a discussion of such connections see Nunberg *et al.* 1994 and Strässler 1982). The chapter on the semantic structure of idioms provides detailed analyses at the individual-example level. This lays the foundation for all the other analytical chapters in this work and, therefore, this chapter is placed first among them.

1.9 Conventions used in this book

This study adopts some conventions in presenting its data and analyses. The most important of these are:

- Multiword unit English translations of Arabic idioms have been given whenever convenient. Otherwise, paraphrastic translations have been provided.
- The translations given are not always exclusive. Sometimes, they have been chosen primarily as conveniently close equivalents. In many cases, these have been obtained from Wehr (1994). However, it should be noted that Wehr's work does not include all the idioms in the data, and in some cases it includes the idiom but does not list all the meanings that have been attested for it in the corpus data.
- Literal translations of Arabic idioms are sometimes given in the line following the gloss, preceded by the abbreviation "lit". I have adopted this approach when I have thought that it would make clearer how the idiomatic components are connected grammatically and semantically at the literal level.
- Proposed English translations of the corpus examples have not been given systematically. Sometimes I have not provided them when I have thought that

the glosses (and the literal translations) suffice to clarify the point under examination.

- Generally, in the case of polysemy, when the semantics of the idiom is not crucial for clarifying the point being discussed, only its most prominent meaning in the corpus has been given in the line following the gloss, leaving more in-depth investigations of the semantics of the expression to be dealt with in other appropriate sections.
- Typically, the base forms of verbal idioms are given in the past tense, singular, masculine, and third person form (see Chapter 2). For the sake of simplicity, the glosses of these base forms have not in general incorporated all these categories of morphological information.
- Very often, the base forms of the idioms have been given before their corpus examples.
- Although the glosses included are usually detailed, occasionally some details that are irrelevant to the point under discussion have been left unmentioned.
- Generally, when parts of the idiom or the corpus example are given in the analytical or explanatory comments to discuss their specific roles, they have been provided in the nominative case if they are nominals, and in the indicative mood if they are verbs that occur in the example in their present tense form.
- Although the terms verb-subject, verbal, nominal, prepositional, adjectival, and adverbial idioms have been used throughout this work, for convenience, the labels V-SBJ, VP, NP, PP, AP, and ADVP idioms have often been used inside the tables that are provided in the coming chapters.
- Sometimes in the analytical chapters, the results of the data analysis and their discussion have been presented separately in different sections. In order to save the reader a lot of referring backwards between the discussion and the examples provided in the analysis, as appropriate, I have provided literal translations of the base forms of the relevant examples in the discussion sections.

2 Methodology and data

This chapter sets out the methodology of the study and provides lists of the idioms investigated along with information on their frequency in the main corpus I have used. It discusses the procedures followed in collecting and classifying Arabic idioms and in selecting the sample for the purpose of investigation. It also gives details on the sources of corpus data, describes the searching procedures, discusses how the corpus data have been prepared for examination, and addresses the status of the language of Arabic newspapers in MSA.

2.1 Collecting Arabic idioms

In order to collect the data needed for this work, idioms have been gathered from some published dictionaries and examples I have observed in everyday interactions and readings.* One important work I have used as a source of Arabic idioms is Sieny *et al.* (1996). This dictionary is interested in Arabic multiword units. However, many of its entries are devoted to types of expression that are not considered idioms according to the characterization adopted in the present study. These include, for instance, literal restricted collocations e.g. in (1), semi-literal restricted collocations e.g. in (2), prepositional verbs e.g. in (3), and proverbs e.g. in (4). Therefore, such examples were excluded.

- | | | | | |
|-----|---|---|-----------------------------------|----------------------------|
| (1) | al-kamm-u
DEF-quantity-NOM
quantity | wa-l-kayf-u
and-DEF-quality-GEN
and quality | | |
| (2) | ghayyara
change.PST
he changed | naghamat-a-hu
tone.F-ACC-M.3SG.POSS
the general character of his speech | | |
| (3) | 'awwala
rely.PST | 'alā
on | | |
| (4) | lā
NEG | ta'dam-u
be.lacking.PRS.F-IND | l-ḥasnā'-u
DEF-beautiful.F-NOM | dhām-a-n
fault-ACC-INDF |
- no one is perfect

The expressions that have been classified as idioms in the step described above, i.e. 680 examples, have been searched for in the *All Newspapers* section in *Arabi-corpus*, which constitutes the major source of corpus data for this study (see section 2.4.1). Only the idioms that occur in this section have been included in the initial list of examples. The other examples that do not occur in this section, 26 idioms, have been excluded. This step has been taken in order to ensure that the idioms collected are in use and that corpus data that can help examine their linguistic behavior are obtainable.

In searching for the occurrences of these idioms in the corpus, I have used search terms that are as loose and inclusive as possible. This has been motivated by a wish to increase the chances of locating any occurrence of the idioms in the corpus texts. However, it is important to note that the searches conducted at this stage were limited, since it was not feasible to look for the incidence of all the 680 idioms in the same detailed way the study uses with its sample, as will be discussed later. Therefore, it might be the case that a few idioms that were excluded actually occur in the corpus texts but were not retrieved during this stage.

Since the study is mainly concerned with determining the different types of idiom and the patterns of idiom behavior in Arabic, the frequency of the examples in the corpus has not been used as a criterion for their inclusion in this initial list of data. At least in principle, an infrequent idiom may represent e.g. a unique structural type or exhibit a pattern of linguistic behavior that may not be shown by other more frequent idioms. Therefore, any idiom that occurs in the corpus even once or twice has been included in the initial list of data.

It might be argued that the infrequency of the expression in the corpus could bring its institutionalization into question. That is, it could be only a non-institutionalized, one-off, figurative expression. In order to ensure that all the examples included in the data are institutionalized, candidate expressions that occur fewer than ten times in the corpus and that are not obtained from Sieny *et al.* (1996) have been included in the initial list of idioms only if they occur in other data sources (see section 2.4.2). The number ten is not significant in itself, but a reasonable criterion was needed.

Following the line of thought in Moon (1998), Riehemann (2001), and Langlotz (2006), broadly synonymous expressions with common lexis have been considered as variants of the same idiom and therefore have been recorded only once in the idiom list. As Moon (1998: 122) puts it:

[t]his view allows newly encountered variant forms to be reconciled with those forms already found, providing further evidence of instability, rather than enforcing either their categorization as completely new items or else their dismissal as deviant.

Another relevant phenomenon is that sometimes more or less the same figurative image manifests itself in a number of expressions that belong to different syntactic classes and have related but different meanings. For example, the following two expressions occur frequently in the corpus data:

If the idiom does not occur in this dictionary but has been encountered in daily interactions and readings, the following rules have been applied as strictly as possible:

- When the idiom contains a verb and its subject, it has been listed in the VS order, with the verb in the past tense form, and the subject in the singular form, unless the corpus data suggest that the expression cannot be used in this way.
- Verbal idioms that include complements other than the subject have been listed in the past tense, third person, masculine, and singular form, as long as they occur in this form in the corpus data. However, there are some cases in which such idioms occur only in the present tense form. In these cases, the idioms have been listed in the present tense form.
- When the idiom shows some lexis substitutability, some effort has been made to determine its most frequent variant in order to use it in its base form. However, the large size of the corpus I use here has made it difficult to ensure that the most frequent variant has been always determined accurately at this early stage.

2.2 Classifying Arabic idioms

Because of the time constraints on the present study, it has not been feasible to examine the instances of this huge number of idioms in the corpus texts. Therefore, I have decided to select a sample of the idioms collected for the purposes of investigation.

In order to ensure a balanced representation of the different types of idiom in the data, two steps have been taken. First, the 654 idioms collected have been categorized according to their syntactic class. As a result, six idiom classes have been identified in the data. Table 2.1 shows the syntactic classes identified and the distribution of the idioms collected over them.

Verb-subject idioms refer to examples that consist of at least one verb and (the syntactic head of) its subject with or without any other constituents. Thus, the minimum realization of this type consists of a verb and a noun that functions as (the head of) its subject. This term should therefore be taken only as a label rather than as a detailed description by itself. The labels *verbal*, *nominal*, *prepositional*, *adjectival*, and *adverbial idioms* refer simply to expressions that are syntactically headed by verbs, nouns, prepositions, adjectives, and adverbs, respectively.

Table 2.1 Syntactic classes of Arabic idioms

<i>Syntactic class</i>	<i>No.</i>
Verb-subject idioms	58
Verbal idioms	304
Nominal idioms	162
Prepositional idioms	92
Adjectival idioms	31
Adverbial idioms	7

Two points are in order. First, I have distinguished verb-subject idioms from verbal ones because of the important syntactic differences between the two types. In particular, Arabic grammar requires that the verb and the subject of the sentence agree in gender, number, and person, according to certain rules. This agreement is not required between the verb and its object. This distinction may therefore facilitate the discussion of the relevant phenomena in the following chapters. Second, it should be noted that the adverbial function in Arabic is very often carried out by nouns and adjectives used in the accusative case (see Ryding 2005 and Badawi *et al.* 2004). Indeed, all the seven adverbial idioms that I have found are syntactically headed by nouns in the accusative.

In the second step, the examples belonging to each of these syntactic categories have been arranged alphabetically within a separate list. In this order, only the idiomatic components have been considered. Elements that the context supplies such as negatives and open-slot fillers have been disregarded when arranging the examples. The actual ordering has been done simply by using the “sort” function in Microsoft Word.

2.3 Selecting the sample

The sample investigated consists of 10 percent of the idioms belonging to each of the syntactic categories in the data. Two major steps have been taken to select this sample from the lists. First, I have calculated how many idioms in each list are going to be examined. Here, as a rule, I have rounded up the fractions to the nearest higher whole numbers. This step applies to all the lists except for the adverbial idiom one, since it contains only seven examples. I have therefore decided to examine two adverbial idioms. Accordingly, the sample contains 6 verb-subject idioms, 31 verbal idioms, 17 nominal idioms, 10 prepositional idioms, 4 adjectival idioms, and 2 adverbial idioms, that is, 70 examples. Second, the actual idioms have been chosen using software that generates random numbers within ranges that can be set by the user.

The random selection of 10 percent of the examples would hopefully reflect all the major phenomena in the initial list of Arabic idioms for two reasons. First, selecting the sample on the basis of the syntactic categorization of idioms ensures that all the syntactic types of the data are covered. Second, arranging the examples alphabetically within separate lists bases the selection process on a criterion that is not on the whole intrinsic to the phenomenon under investigation and, therefore, is not likely to bias the selection process and, in turn, the results of the analysis in any major way.

On occasion, I have discussed other examples that are in the initial list of idioms but that have not been selected in the sample. These have been only included in particular sections when they show strikingly different linguistic behavior compared with the behavior of the sample idioms with regard to the point under discussion. Such idioms have been referred to appropriately in text, and they have not been considered when I have given any statistical information.

2.4 Corpus data: sources, collection procedures, and preparation

2.4.1 *Arabicorpus*

The study derives its data mainly from the *All Newspapers* section of *Arabicorpus*. This is a corpus of Arabic that is accessible through the web[†] and that has been under ongoing development by Dilworth Parkinson at Brigham Young University. In the following chapters, unless otherwise stated, when the label *Arabicorpus* is used, it should be taken to refer solely to this section.

Arabicorpus is a large corpus that makes use of several powerful searching capabilities. In the following two sections, I provide more information on this corpus and discuss why I have decided to rely mainly on the *All Newspapers* section of it. The information provided below on the contents, structure, and searching capabilities of *Arabicorpus* reflects its state at the time when the major stage of data collection for this study came to an end in September 2007.

2.4.1.1 *Contents and structure of Arabicorpus*

The *All Newspapers* section includes raw texts, i.e. texts which are not annotated grammatically or semantically, that come from five major Arabic newspapers. Its total word count is 83,519,701. These newspapers include one year of *Al-Ahram* (1999, Egypt), two years of *Al-Hayat* (1996 and 1997, London – Saudi Arabia – Lebanon) in two separate sub-corpora, a half year of *Al-Tajdid* (2002, Morocco), a half year of *Al-Watan* (2002, Kuwait), and approximately a year of *Al-Thawra* (after 2000, Syria). All these can be searched separately or together.

This section does not include the whole texts of the printed papers. For example, it does not contain classifieds and newspaper advertisements. Rather, it only includes, in the developer's words, "the serious parts which ended up on the website" of each newspaper (Dilworth Parkinson, email communication, 21 June 2009).

The word total for each newspaper in this section is given in Table 2.2. The term word is understood here as a sequence of alphanumeric characters that is bounded on either side by spaces, punctuation, or nothing, i.e. at the beginning and end of paragraphs (Dilworth Parkinson, email communication, 19 June 2009).

Table 2.2 The word count of each newspaper in *Arabicorpus*

<i>Newspaper</i>	<i>Word count</i>
Al-Ahram, 1999	16,475,979
Al-Hayat, 1996	21,564,239
Al-Hayat, 1997	19,473,315
Al-Tajdid, 2002	2,919,782
Al-Watan, 2002	6,454,411
Al-Thawra, after 2000	16,631,975

Besides this section, *Arabicorpus* contains several other sub-corpora. Table 2.3 refers to these and gives information on their word counts.

The Penn Treebank news data is not included in the *All Newspapers* category. The pre-modern corpus includes The Quran, 1001 Nights, and some works of medieval philosophy and medicine. The modern literature corpus includes a number of Arabic novels, some plays, and some short stories. The non-fiction category contains some literary criticism works, other scholarly works, some political speeches, and a number of official UN and other diplomatic documents. The Egyptian Arabic category includes some literary works that contain an amount of colloquial Arabic and some material from the Egypt Chat website. Each of these components can be searched separately or as part of *Arabicorpus* as a whole. Other sub-corpora group and regroup in various configurations.

Since this work is only concerned with idioms in MSA, the pre-modern and colloquial texts are not of concern here. The decision to rely mainly on the *All Newspapers* section rather than the other categories that relate to MSA, i.e. modern literature, non-fiction, and Penn Treebank news data, has been taken for a number of practical reasons.

First, the newspaper category is by far the largest section in *Arabicorpus*. Combined together, the other three MSA-related categories contain 2,180,034 words, as opposed to more than 83 million words in the *All Newspapers* section. Therefore, in terms of size, the latter is a much better candidate for revealing the major patterns relevant to the data.

Second, and very importantly, it is not possible to combine in performing the corpus searches one or more of these three categories with the newspapers. Therefore, in order to search these rather small sub-corpora, one should repeat the whole set of searches performed for each idiom. This does not seem feasible due to the complications involved in the searching procedures (see section 2.4.1.2 for details).

Third, I have done some limited searches of these three categories but I have often retrieved only few or even no relevant results. Two reasons may be proposed for this: first, the rather small size of these sections; second, it might be argued that, particularly in the case of modern literature, the authors, in their attempt to be original and genuinely creative, may refrain from deploying some types of

Table 2.3 Components of *Arabicorpus* other than newspapers

<i>Corpus</i>	<i>Word count</i>
The Quran	84,532
Novels	928,776
Modern Literature	1,001,899
Pre-modern Texts	912,996
Non-fiction	579,545
Penn Treebank News Data	598,590
Medieval Science	223,249
1001 Nights	557,908
Egyptian Colloquial Arabic	157,099

prefabricated figurative multiword units, especially those that might be considered trite on account of overuse.

Finally, to varying degrees, the topics of the texts in these three sub-corpora are also represented in the *All Newspapers* category, since this includes texts that belong to many different domains, e.g. news articles, sports, religion, literary works such as short stories, science, technology, literary criticism, philosophy, and social, political, and economic analyses (also see below the discussion of the status of newspaper Arabic, and Moon's (1998) comment regarding the use of idioms in newspaper language in general).

Having said that, I have sometimes consulted these three categories in order to establish, or not, a specific use of an idiom. A case in point is where an idiom does not occur in the *All Newspapers* section within a certain grammatical structure that has been connected in the linguistic literature on Arabic with the modern literary style. In such a case, I have consulted the modern literature section of *Arabicorpus* so as to determine whether the idiom is used or not in this structure in its texts, since this knowledge may assist in deciding whether this use is acceptable in certain writing styles or that it does not seem to be employed in modern Arabic writings.

2.4.1.2 *Searching Arabicorpus*

Arabicorpus has its own searching algorithm which utilizes many of the advantages that regular expressions can provide. In computing, these expressions are strings that can be written using a formal language to set a search pattern within an electronically stored text. The results retrieved from the corpus consist mainly of the search term with up to ten words before and after it in context. The corpus also gives access to the whole texts where the idioms occur. This was very useful during the data analysis stage, because, in many a case, I needed to read more parts of the text so as to establish or not a certain use of the idiom, particularly on the semantic and discursive levels.

In the following, I discuss some of the complications involved in locating the occurrences of the data examples in the corpus texts and also provide details on how I have conducted my searches.

Arabicorpus does not enable its user to easily locate the co-occurrence of two or more words within ranges that they can set. Thus, in order to find the examples of an idiom consisting, for instance, of two words, such as a verb and its object, without any intervening elements between them, a string of the two words written adjacent to each other must be used as a search term. In order to search for the same two words with only one intervening element between them, another search term is needed where regular expressions are deployed to allow for any single sequence of alphanumeric characters to occur between the idiomatic elements. A third search with a different search term is still needed to look for the same idiom with two intervening elements between its components, and so on.

In many cases, such searches have been performed six times for each idiom in order to retrieve the examples of the idiom in the corpus with up to five intervening

words between its two components. For example, in order to locate the occurrences of the idiom:

<i>shakhṣi/shay'</i>	ṭawwaqa	‘unuq-a	<i>shakhṣi</i>
<i>sb_i/sth</i>	surround.PST	neck-ACC	<i>sb_j</i>

lit. *sb_i/sth* surrounded *sb_j*'s neck
sb_i or *sth* (e.g. trust) greatly honored *sb_j*, or *sth* (e.g. debt) burdened *sb_j*

the following six search terms have been exploited:

```
Tw\w*q\w* \w*cn\w*q
Tw\w*q\w* \w+ \w*cn\w*q
Tw\w*q\w* \w+ \w+ \w*cn\w*q
Tw\w*q\w* \w+ \w+ \w+ \w*cn\w*q
Tw\w*q\w* \w+ \w+ \w+ \w+ \w*cn\w*q
Tw\w*q\w* \w+ \w+ \w+ \w+ \w+ \w*cn\w*q
```

These search terms are based on the transliteration scheme adopted by *Arabi-corpus*.[‡] This system generally ignores the symbols for short vowels. Also, in this particular example, two other differences between this scheme and the one adopted by the present study can be pointed to, i.e. the T and c above stand respectively for the ṭ and ‘ in the system I use.

The following regular expressions are utilized in the example: \w+, which stands for one or more word characters in a row, and \w*, which stands for none or more word characters in a row. The intensive use of regular expressions at the beginning, middle, and end of the idiomatic words has successfully captured the variations that could occur as a result of different derivational and inflectional processes.

In general, this work follows the foregoing searching procedure when it is, in principle, possible according to the general grammatical rules of Arabic to have any number of intervening elements between the idiomatic components. This is the case, for example, in verbal idioms that contain, in addition to the verb, an object and/or a prepositional adjunct phrase. These types of idiom, as will be seen later, are frequent in the data.

In order to put these complications into perspective, one may quote Langlotz (2006) describing how he searched the British National Corpus (BNC) for the occurrences of the idioms in his data:

[T]he corpus was searched with the client programme *SARA-32 version 0.98*, the customary search tool supplied with the *BNC World Edition* [. . .] To find tokens of a given idiom such as *rock the boat*, I used the *Query Builder*, a general search mode customised by means of a visual interface. In general, I specified the searches by asking the program to look for combined instances of the lexical key-constituents, i.e. *rock* and *boat*, within a span of up to 10 words. To include all morphological inflections of these head-words, I used the *Lancaster lemma-scheme*.

(Langlotz 2006: 228, his italics)

Two points may be highlighted here: first, the relative ease of locating the co-occurrences of the idiomatic components within ranges that can be set by the researcher and, second, having access to an incorporated morphological analyzer that enables the user to retrieve all the different morphological inflections simultaneously.

Due to the properties of *Arabicorpus* described above, in many a case, I needed to conduct more than ten corpus searches, in order to retrieve the examples of one and the same idiom.

In general, when the idiom contains a word that is not frequent in use, I have only used this word as a search term. This has increased the chances of retrieving all the instances of the relevant idiom in the corpus. However, this approach obviously cannot retrieve any corpus examples where this word is omitted or replaced with any other lexical item(s). As Moon (1998) puts it, locating examples of lexical substitutions has been “a matter of serendipity” (p. 51).

Certain formal variations in idioms pose certain difficulties in locating their incidence in the corpus. For example, lexical substitutions are particularly problematic when the idiom consists only of words that are frequent in use, since using only one of such words as a search term would retrieve thousands of irrelevant hits which would be practically impossible to handle. Therefore, the search terms in such cases have had to consist of at least two idiomatic words and to be as loose and inclusive as possible in order to be able to capture as many relevant corpus examples as possible.

Furthermore, in some cases, the general grammar rules allow, at least in principle, for the idiomatic words to change their positions relative to each other while preserving their typical grammatical relationships, e.g. when the idiom consists of a verb and its subject or when it contains two or more conjuncts. Also, in many cases of embedding (see Chapter 5), the order in which the idiomatic words occur differs from that of the base form. So as to capture such uses in *Arabicorpus*, there are, in theory, two options. The first is to conduct the corpus searches using search terms that utilize regular expressions that can locate the two (or more) idiomatic components in any possible order. The other option is to conduct several sets of searches, changing the order of the idiomatic components in each.

Practically, the first option, which often involves the use of very complicated search terms, was problematic. In many cases, the searching process took a very long time such that the web browsers I use failed to retrieve anything from *Arabicorpus* and error messages occurred. Therefore, I have chosen, or even have been forced to choose, the second way, since I have found it, although tedious, less time-consuming.

2.4.2 Additional data sources

Although the judgments made in this study are chiefly based on the relevant data in *Arabicorpus*, certain of these judgments are also based on additional data obtained from other sources. These sources include some investigations of a number of Arabic websites and of The Linguistic Data Consortium Arabic News

Text corpus (for simplicity, I will use *LDC Arabic Corpus* to refer to this work. However, it should be borne in mind that the LDC has developed other Arabic corpora in addition to this one). This approach has been adopted in order not to undermine relevant phenomena that do not occur in *Arabic corpus*. It has proven particularly useful when the idiom does not occur frequently in its texts.

As for the former, I have occasionally examined examples obtained from some Arabic websites that are written in MSA. These are mostly, but not exclusively, online Arabic newspapers. In these cases, in order to locate relevant examples of the idiom under examination, I have simply searched for its occurrences using the Google search engine.

As for the latter, the *LDC Arabic Corpus* contains texts from two news agencies and three newspapers. The news agencies are Agence France Presse and Xinhua News Agency. The newspapers are Al-Hayat, Al-Nahar, and Al-Ummah. I have not conducted systematic, exhaustive examinations of this corpus because it is extremely large such that it cannot be investigated thoroughly by an individual researcher. This corpus consists of 489,729,594 words. In addition, although the *LDC Arabic Corpus* makes use of some advanced searching tools such as wildcards, it does not utilize many of the functions of regular expressions.

Lastly, on occasion, I have carried out some informal interviews with native speakers of Arabic. However, it should be stated that asking the informants about how they would use and vary a certain expression is not likely to reveal all the forms acceptable in their intuitions. Simply, they could forget to mention some of these forms. Furthermore, this study is interested in a considerable number of questions, i.e. semantic, discursive, lexical, and grammatical ones. It was not therefore viable to ask the informants all the relevant questions about all the examples in the data, since people could just lose interest in such a task. For these reasons, I have conducted my interviews only when convenient. They have been particularly useful in shedding light on how native speakers perceive the relationship between the literal and idiomatic meanings of some of the examples in the data.

My informants are adult native speakers of Arabic who have obtained at least a first university degree in a field other than Arabic studies from a university in an Arabic-speaking country. First, adult speakers normally possess well-developed language systems that are not typically available to children. Possessing such developed systems appears to be of particular significance in the context of a study whose focus is non-literal, idiomatic language and the variation in its uses. Second, MSA is a variety of Arabic that is learnt almost exclusively through formal instruction at schools in Arab countries. I have, however, excluded native speakers who hold degrees in Arabic studies, as they may not be a faithful representation of the average native Arabic speaker today, due to their specialist expertise in language.

Such a combination of different data sources, sometimes referred to as *triangulation* (Perry 2005), would hopefully reveal the major phenomena in the linguistic behavior of Arabic idioms and would help validate the interpretations presented in the coming chapters.

2.4.3 *Preparing the corpus data*

The concordance lines retrieved from *Arabicorpus* have been stored in Microsoft Excel files or as HTML files when more convenient. This gave me the means to conduct many analyses without having to have instant access to the corpus web page all the time. As noted above, the concordance lines consist only of the search term with up to ten words before and after it in context. Therefore, when knowledge of more portions of the texts was needed during the analysis stage in order to establish one or another use of an idiom, these files were not of much help, and I had to consult the full texts through the corpus web page itself.

Arabic idioms differ with regard to their frequency in *Arabicorpus*. While some idioms only occur a few times, others occur very often, sometimes several thousand times. Two points need to be noted here. First, because the rare occurrences of an idiom could undermine some phenomena in its linguistic behavior, other sources of data have often been consulted when the idiom occurs only a few times in *Arabicorpus* texts.

Second, it has not been possible to examine all the occurrences of very frequent idioms. In order to deal with a similar problem, Langlotz (2006) restricted his investigations of any idiom that occurs more than 150 times in the BNC to only a randomly selected 100 examples (p. 228). In this study, I have decided to limit my investigations to a random selection of 200 concordance lines for any idiom that occurs more than this number of times in the corpus. I have decided to include more examples than Langlotz did mainly because the corpus I use is relatively smaller than the BNC, which contains 100 million words.

Three final points need to be noted in relation to the preparation of the corpus data. First, the study is interested only in the non-literal uses of the data; therefore cases where the idioms are used literally or as similes have not been considered when giving any statistical information in this work. However, it should be mentioned that these uses have sometimes made clear how native speakers of Arabic perceive the idioms in question. Second, I have excluded the very few cases where the idioms themselves are the subject of some linguistic analysis. Finally, sometimes, the searching algorithm retrieved the same corpus example two or more times. In such cases, the example has been counted only once.

2.5 **The language of Arabic newspapers**

According to Ryding (2005: 8):

[a] fully agreed-upon definition of MSA does not yet exist, but there is a general consensus that modern Arabic writing in all its forms constitutes the basis of the identity of the language. Modern writing, however, covers an extensive range of discourse styles and genres ranging from complex and conservative to innovative and experimental. Finding a standard that is delimited and describable within this great range is a difficult task; however, there is an identifiable segment of the modern Arabic written language used for

media purposes, and it has been the focus of linguists' attention for a number of years because of its stability, its pervasiveness, and its ability to serve as a model of contemporary written usage. Dissemination of a written (and broadcast) prestige standard by the news media is a widespread phenomenon, especially in multilingual, diglossic, and multi-dialectal societies.

In actual fact, some researchers have defined MSA as the language of Arabic news media (Monteil, 1960, cited in Ryding 2005: 8).

Several recent works on MSA have derived much or all of their data from newspaper Arabic, e.g. Abdel-Aziz (2003), Badawi *et al.* (2004), and Ryding (2005). In the following extracts, Ryding (2005) defends her use of media Arabic as the major source of data for her comprehensive work on MSA grammar and argues for the significance of the newspaper language for the learners of Arabic as a foreign language:

Media Arabic was chosen as a main source of data for this text because of its contemporaneousness, its coverage of many different topics, and the extemporaneous nature of daily reporting and editing. As a primary source of information about and from the Arab world, newspaper and magazine language reflects Arab editorial and public opinion and topics of current interest. Various subject matter and texts were covered, ranging from interviews, book reviews, feature stories, religion and culture, and sports reports, to straight news reports and editorials. (p. xix)

There are doubtless those who would assert that the ordinariness of media language causes it to lack the beauty and expressiveness of literary Arabic, and therefore that it is unrepresentative of the great cultural and literary achievements of the Arabs. To those I would reply that the very ordinariness of this type of language is what makes it valuable to learners because it represents a widely used and understood standard of written expression. As Owens and Bani-Yasin (1987, 736) note, "the average Arab is probably more exposed to this style than to most others, such as academic or literary writing." In fact, it is a vital and emergent form of written language, being created and recreated on a daily basis, covering issues from the mundane to the extraordinary. With limited time to prepare its presentation style, media Arabic reflects more closely than other forms of the written language the strategies and structures of spontaneous expression. (p. xix)

Holes (2004) also notes that "constant exposure to this 'media MSA' seems to be having far-reaching effects on the vocabulary, grammar, and phraseology of the Arabic used by educated Arabs in many other contexts, written or spoken" (p. 314).

As the major style of MSA that the average Arab is exposed to, media Arabic thus affects how educated Arabs use their language (creatively) and influences their judgments about what is acceptable or unacceptable in terms of language

use. Also, the fact that newspaper language closely reflects the strategies and structures of spontaneous expression is of great importance in the context of a study on idioms, for these, and formulaic expressions in general, are often associated with spontaneous ways of verbal communication (see Carter 2004 and Cameron 2001). Moon (1998: 121) also notes that:

while variations [in multiword units] occur across the range of text types, it is [*sic*] often associated with journalism. Variations found in journalism cannot be dismissed out of hand as mannerism and journalese. In fact, journalism represents the cutting edge of language change, or the popularization of language change: variations fossilizing here may foreshadow what will later become [*sic*] institutionalized more widely.

In addition, the fact that newspaper Arabic is busily involved in the ongoing intensive translation process particularly from English attaches more importance to it in the context of this study. As noted in the previous chapter, Arabic has expanded its phraseological repertoire with many loan translations from English. Using newspaper Arabic as the major source of data in a study on idioms may therefore provide the means to see this process of idiom creation or re-creation in action.

In the light of the foregoing observations, one may say that using newspaper Arabic (or even newspaper language in general) as the major data source for a study that is concerned with idioms in a variety that predominates in written communication can be regarded as an instance of employing *information-rich samples*. Such samples are exploited in research because they contain many examples of the phenomenon under examination and, therefore, can yield considerable insight into it (see Perry 2005).

The central role of newspaper language in present-day linguistic research on MSA in general can also be seen in the availability of several corpora of different sizes that are based on texts obtained (mainly) from Arabic newspapers. For example, it has already been noted that the newspaper section constitutes the largest section of *Arabic corpus*. In addition, the LDC has developed several corpora based on texts derived from Arabic newspapers. Finally, Al-Sulaiti (2004) has developed a corpus of contemporary Arabic in which she relies heavily on texts obtained from online Arabic magazines and newspapers (for more examples of available corpora of MSA, see Al-Sulaiti 2004).

Nevertheless, there is another factor that cannot be ruled out when accounting for the availability of such corpora of newspaper Arabic. That is, many Arabic newspapers are accessible in machine-readable form, which makes the process of compiling an electronic corpus smoother.

2.6 Data examples and their frequency

In the following tables, I list the sample idioms according to their syntactic classes. In each table, the examples are sequenced in accordance with their alphabetical order in Arabic. Here, only the idiomatic components are considered, that is, open

slot fillers and negatives are not. Each idiom is accompanied by information on its frequency in the *All Newspapers* section of *Arabicorpus*. In addition, comments, particularly on the translations, are sometimes given.

A number of preliminary points are in order. First, in general, the meanings given for the examples below are their most prominent ones in the corpus. Sometimes, other less salient meanings exist. For simplicity, these are not provided here. Second, although I have done my best to give accurate information on the frequency of the examples, it might be the case that some of their corpus occurrences have not been retrieved due to one or another of the difficulties depicted above or due to the typographical errors that sometimes occur in the corpus texts.

In this context, it was particularly difficult to determine accurately the incidence of some very frequent expressions by reading through their concordance lines one by one to exclude, for instance, literal uses and any duplicates. Therefore, in such cases, the numbers provided reflect mainly how many concordance lines were retrieved for each expression. However, it should be noted that these sometimes include instances of literal uses and/or duplicates. These cases have been labeled in the following tables with asterisks that are provided next to the relevant numbers.

Finally, in what follows, whenever convenient, English multiword unit translations have been provided. When the English expression is only a semi-equivalent of or seems to be stylistically different from its Arabic counterpart, it has been included in brackets after a paraphrastic translation, for the purposes of comparison. However, it should be mentioned that such semi-equivalent or stylistically different translations have been presented mainly when the English expression is based on the same figurative image that underlies the Arabic idiom or on one that is similar to it. Also, while proposing English translations is, to a degree, sustained practice throughout this work, I provide these semi-equivalents only in the following tables.

Table 2.4 Verb-subject idioms

No.	Idiom	F.
1	<i>shay'</i> 'akhadha ba'ḍ-u-hu bi-riqāb-i ba'ḍ <i>sth</i> hold.on.PST 8 some-NOM-M.3SG.POSS to-neck.PL-GEN some lit. the parts of <i>sth</i> held on to each other's necks <i>sth</i> (e.g. a set of ideas, texts, or problems) interconnected	
2	'afala najm-u <i>shakhṣ/shay'</i> 74 set.PST star-NOM <i>sb/sth</i> lit. the star of <i>sb/sth</i> set the glory or fame of <i>sb/sth</i> ended, (cf. the sun sets on <i>sth</i>)	
3	inqabaḍa ṣadr-u <i>shakhṣ</i> 14 contract.PST chest-NOM <i>sb</i> lit. <i>sb</i> 's chest contracted <i>sb</i> felt depressed	

(Continued Overleaf)

Table 2.4 Continued

No.	Idiom	F.
4	tajammada l-dam-u fi 'urūq-i shakhs freeze.PST DEF-blood-NOM in vein.PL-GEN sb the blood froze in sb's veins Also, tajammada l-dam-u fi 'urūq-i shay' solidify.PST DEF-blood-NOM in vein.PL-GEN sth sth (e.g. an organization) became completely inactive	8
5	lā taqūm-u li-shakhs/shay' qā'imat-u-n NEG stand.up.PRS.F-IND for-sb/sth pillar.F-NOM-INDF lit. none of sb/sth's pillars stands up sb/sth is completely ruined (e.g. socially or professionally)	3
6	dāra-t rahā shay' spin.PST-F quern.F sth lit. the quern of sth spun sth (e.g. a war, elections, or disagreement) took place in a violent way	122

Table 2.5 Verbal idioms

No.	Idiom	F.
1	'adāra l-khadd-a l-'ākhar-a turn.PST DEF-cheek-ACC DEF-other-ACC he turned the other cheek	28
2	'arghā wa-'azbada foam.PST and-froth.PST he fumed with rage (cf. <i>froth at the mouth</i>)	27
3	iṣṭāda fi l-mā'-i l-'akir-i fish.PST in DEF-water-GEN DEF-cloudy-GEN he fished in troubled waters	161
4	'akmala niṣf-a dīn-i-hi perfect.PST half-ACC religion-GEN-M.3SG.POSS lit. he perfected the [other] half of his religion he got married	4
5	'amsaka l-'aṣā min-a l-muntaṣaf-i hold.PST DEF-stick at DEF-middle-GEN lit. he held the stick at its middle point he struck a balance	40
6	bala'a l-ṭu'm-a swallow.PST DEF-bait-ACC he swallowed the bait	10
7	taqallaba 'alā jamr-i shay' turn.over.PST on ember.PL-GEN sth lit. he turned over the embers of sth he severely suffered from sth	7
8	hāraḇa ṭawāhīn-a l-hawā'-i fight.PST mill.PL-ACC DEF-air-GEN lit. he fought the air mills he tilted at windmills	44

9	ḥafara qabr-a shakhṣ/shay' bi-yad-i-hi	46
	dig.PST grave-ACC sb/sth with-hand-GEN-M.3SG.POSS	
	lit. he dug the grave of <i>sb/sth</i> with his own hand	
	he put an end to e.g. the success or career of <i>sb</i> , or he put an end to <i>sth</i> , (cf. <i>dig one's own grave</i>)	
10	dassa l-summ-a fi l-'asal-i	50
	put.secretly.PST DEF-poison-ACC in DEF-honey-GEN	
	lit. he secretly put the poison in the honey	
	he deceptively promoted something bad (e.g. a way of thinking or behaving) by presenting it in an advantageous or attractive way or presenting it within something else good	
11	dhāqa marārat-a shay'	167
	taste.PST bitterness-ACC sth	
	lit. he tasted the bitterness of <i>sth</i>	
	he experienced the suffering of <i>sth</i>	
12	rafa'a rāyat-a shay'	620*
	raise.PST flag-ACC sth	
	lit. he raised the flag of <i>sth</i>	
	he supported <i>sth</i> , he made <i>sth</i> triumphant, or he exhibited <i>sth</i> (usually a pattern of behavior such as disobedience), (cf. <i>wave the flag of sth</i> , <i>wave the banner of sth</i>)	
13	shakhṣ/shay' sābaqa l-rīḥ-a	15
	sb/sth race.PST DEF-wind-ACC	
	<i>sb/sth</i> ran very fast, <i>sb</i> acted very fast, or <i>sth</i> (e.g. the performance of a company) progressed very quickly	
14	saraḥa wa-mariḥa	99
	move.freely.PST and-enjoy.oneself.PST	
	lit. he moved freely and enjoyed himself	
	he did as he liked	
15	shadda l-ḥizām-a	69
	tighten.PST DEF-belt-ACC	
	lit. he tightened the belt	
	he tightened his belt	
16	ṣabba l-mā'-a l-bārid-a 'alā shakhṣ/shay'	12
	pour.PST DEF-water-ACC DEF-cold-ACC on sb/sth	
	he poured cold water on <i>sth</i> (e.g. a proposal), or he did something that reduced the tension of <i>sb/sth</i>	
17	ḍaraba 'uṣfūr-ayni bi-ḥajar-i-n	130
	hit.PST sparrow-DU.ACC.INDF with-stone-GEN-INDF	
	he killed two birds with one stone	
18	shakhṣ _i /shay' ṭawwaqa 'unuq-a shakhṣ _j	16
	sb _i /sth surround.PST neck-ACC sb _j	
	lit. <i>sb_i/sth</i> surrounded <i>sb_j</i> 's neck	
	<i>sb_i</i> or <i>sth</i> (e.g. trust) greatly honored <i>sb_j</i> , or <i>sth</i> (e.g. debt) burdened <i>sb_j</i>	
19	'affara jabhat-a-hu	4
	cover.with.dust.PST forehead.F-ACC-M.3SG.POSS	
	lit. he covered his forehead with dust	
	he humiliated himself	
20	ghariqa ḥattā 'uthun-ay-hi fi shay'	6
	sink.PST up.to ear-DU.GEN-M.3SG.POSS in sth	

(Continued Overleaf)

Table 2.5 Continued

No.	Idiom	F.
	lit. he sank up to his [two] ears in <i>sth</i> he was up to his ears in <i>sth</i>	
21	<i>shakhṣ/shay'</i> _i fataḥa l-nār-a 'alā <i>shakhṣ/shay'</i> _j <i>sb/sth</i> _i open.PST DEF-fire-ACC at <i>sb/sth</i> _j <i>sb</i> _i verbally attacked <i>sb/sth</i> _j , or <i>sth</i> _i caused <i>sb/sth</i> _j to be criticized (cf. <i>to open fire at sb/sth</i>)	71
22	qalaba zahr-a l-mijann-i li- <i>shakhṣ/shay'</i> turn.around.PST back-ACC DEF-shield-GEN to- <i>sb/sth</i> lit. he turned around the back of the shield towards <i>sb/sth</i> he became hostile to <i>sb/sth</i> after he was a friend or supporter of them/it	25
23	kasara shawkat-a <i>shakhṣ/shay'</i> break.PST thorn.F-ACC <i>sb/sth</i> lit. he broke <i>sb/sth</i> 's thorn he deprived <i>sb/sth</i> of their/its power Another literal meaning of <i>shawkah</i> is 'sting'.	126
24	<i>shakhṣ/shay'</i> _i wa-man/mā laffa <i>sb/sth</i> _i and-REL.HUM/REL.NONHUM turn.round.PST laff-a- <i>hu</i> _i turning.round-ACC-M.3SG.POSS _i lit. <i>sb/sth</i> _i and that who/which turned round in the same way as <i>sb/sth</i> _i did <i>sb</i> _i and that who is of his type, <i>sth</i> _i and that which is of its kind, or <i>sb</i> _i and that who does just as he does (cf. <i>and suchlike</i>)	41
25	<i>shay'</i> marra murūr-a l-kirām-i bi- <i>shakhṣ</i> <i>sth</i> pass.PST passing-ACC DEF-noble.people-GEN by- <i>sb</i> lit. <i>sth</i> (e.g. an event or anniversary) passed as noble people do by <i>sb</i> <i>sth</i> passed unnoticed by <i>sb</i> Also, <i>shakhṣ</i> marra murūr-a l-kirām-i bi- <i>shay'</i> <i>sb</i> pass.PST passing-ACC DEF-noble.people-GEN by- <i>sth</i> lit. <i>sb</i> passed as noble people do by <i>sth</i> <i>sb</i> overlooked <i>sth</i> , or <i>sb</i> treated <i>sth</i> with disdain	190
26	naza'a fatīl-a <i>shay'</i> pull.out.PST fuse-ACC <i>sth</i> lit. he pulled out the fuse of <i>sth</i> he defused <i>sth</i> (e.g. a dangerous situation)	403
27	waḍa'a <i>shakhṣ/shay'</i> 'ala l-raff-i put.PST <i>sb/sth</i> on DEF-shelf-GEN lit. he put <i>sb/sth</i> on the shelf he made <i>sb/sth</i> no longer useful or desirable (cf. <i>on the shelf</i>)	52
28	waqafa fī tarīq-i <i>shakhṣ/shay'</i> stand.PST in way-GEN <i>sb/sth</i> he stood in the way of <i>sb/sth</i>	180
29	yahruth-u fī l-baḥr-i plough.PRS-IND in DEF-sea-GEN lit. he is ploughing in the sea he is ploughing the sand	21

30	<i>lā</i> <i>yūṣaddiq-u</i> ‘ <i>ayn-ay-hi</i> NEG believe.PRS.M.3SG-IND eye-DU.ACC-M.3SG.POSS he does not believe his eyes	31
31	<i>yanfukh-u</i> <i>fī</i> <i>qirbat-i-n</i> <i>maqṭū‘at-i-n</i> blow.PRS-IND into bagpipe-GEN-INDF cut.PPTCP-GEN-INDF lit. he is blowing into a cut-off bagpipe he is trying in vain to convince other people to take certain actions to improve a situation (cf. <i>flog a</i> <i>dead horse</i>)	13

Table 2.6 Nominal idioms

No.	Idiom	F.
1	<i>‘ibrat-u-n</i> <i>fī</i> <i>kawmat-i</i> <i>qashsh-i-n</i> needle-NOM-INDF in stack-GEN hay-GEN-INDF a needle in a haystack	9
2	<i>‘uṣṭuwānat-u-n</i> <i>mashrūkh-at-u-n</i> record.F-NOM-INDF crack.PPTCP-F-NOM-INDF lit. a cracked record repetitious boring talk about a particular topic (cf. <i>a broken record</i>)	40
3	<i>allutayyā</i> <i>wa-llatī</i> REL.F.SG.DIMIN and-REL.F.SG much ado, or lengthy difficult discussions	2
4	<i>baqarat-u-n</i> <i>ḥalūb-u-n</i> cow.F-NOM-INDF milch-NOM-INDF milch cow (cf. also <i>cash cow</i>)	30
5	<i>tuffāḥat-u</i> ‘ <i>ādam</i> apple-NOM Adam Adam’s apple or (esp. in literary language) the forbidden fruit	2
6	<i>ḥibr-u-n</i> ‘ <i>alā</i> <i>waraq-i-n</i> ink-NOM-INDF on paper-GEN-INDF lit. ink on paper inactive (e.g. of decisions or laws) (cf. <i>on paper</i>)	383
7	<i>ḥalqat-u-n</i> <i>mufraqh-at-u-n</i> ring.F-NOM-INDF cast.in.a.mold.PPTCP-F-NOM-INDF lit. a cast-in-a-mold ring vicious circle	100
8	<i>dam-u-n</i> ‘ <i>azraq-u</i> blood-NOM-INDF blue-NOM blue blood	12
9	<i>rimāl-u-n</i> <i>mutaḥarrik-at-u-n</i> sand.PL.F-NOM-INDF move.APTCP-F-NOM-INDF quicksand (i.e. very dangerous situations which are difficult to escape from), or an unreliable basis for <i>sth</i> such as a plan or project	97
10	<i>shadd-u-n</i> <i>wa-jadhb-u-n</i> tugging-NOM-INDF and-pulling-NOM-INDF great tension or struggle (e.g. in a strained political or social state)	123

(Continued Overleaf)

Table 2.6 Continued

No.	Idiom	F.
11	ghuṣn-u l-zaytūn-i branch-NOM DEF-olive-GEN olive branch	49
12	qarṣat-u ʿudhun-i-n pinch-NOM ear-GEN-INDF lit. a pinch of an ear a light non-physical punishment that serves as a warning of a possible cruel one (cf. <i>a rap across the knuckles</i> and <i>a slap on the wrist</i>)	1
13	qimmat-u jabal-i l-jalīd-i top-NOM mountain-GEN DEF-ice-GEN lit. the top of the ice mountain the tip of the iceberg	30
14	marbaṭ-u l-faras-i place.where.sth.is.tied.up-NOM DEF-mare-GEN lit. the place where the mare is tied up the most important point (e.g. in a problem or argument)	56
15	nār-u-n ʿalā ʿalam-i-n fire-NOM-INDF on mountain-GEN-INDF lit. a fire on a mountain very well-known entity	2
16	wāsiṭat-u l-ʿiqd-i middle.part-NOM DEF-necklace-GEN lit. the middle jewel of the necklace the most important or best part of something good (cf. <i>the jewel in the crown</i> and <i>the centerpiece of sth</i>)	39
17	al-yad-u l-ʿulyā DEF-hand.F-NOM DEF-upper.F the upper hand or (in religious context) the giver of charity	79

Table 2.7 Prepositional idioms

No.	Idiom	F.
1	shakṣ/shayʿ _i bi-rummat-i-hi _i sb/sth _i with-worn.out.piece.of.rope-GEN-M.3SG.POSS _i lit. sb/sth _i with their/its worn out lead all sb _i (e.g. a group of people), or the whole sth _i , (also can be used adverbially to mean ‘completely’)	3,000*
2	bayna nār-ayni between fire-DU.GEN-INDF lit. between two fires between the devil and the deep blue sea	38
3	taḥta l-ʿaḏwāʿ-i under DEF-light.PL-GEN in the spotlight	65
4	ʿala l-ṭāwilat-i on DEF-table.F-GEN on the table	900*

5	‘alā ṣaffīh-i-n on sheet.iron-GEN-INDF lit. on hot sheet iron (of a person) very anxious, or (of a situation) very tense or unstable (cf. <i>to be on tenterhooks</i> and <i>to be like a cat on a hot tin roof</i>)	sākhin-i-n hot-GEN-INDF	25
6	<i>shakhṣ/shay’</i> _i <i>sb/sth</i> _i lit. <i>sb/sth</i> _i on their/its father’s young female camel <i>sb/sth</i> _i (e.g. a group of people or things) all together, all <i>sth</i> _i (also can be used adverbially to mean ‘completely’)	on young.camel-F-GEN father-GEN-M.3SG.POSS _i	78
7	‘alā marmā at distance.that.sth.is.thrown lit. within the distance that a stone is thrown of proximity in space or time (cf. <i>at a stone’s throw</i>)	ḥajar-i-n stone-GEN-INDF	31
8	fī ‘a‘qāb-i in heel.PL-GEN lit. in the heels of close behind <i>sb</i> , or immediately after <i>sb/sth</i> (cf. <i>on the heels of</i>)	<i>shakhṣ/shay’</i> <i>sb/sth</i>	4,320*
9	fī daw’-i in light-GEN in the light of	<i>shay’</i> <i>sth</i>	7,500*
10	min warā’-i from behind-GEN behind <i>sb</i> ’s back	zahr-i back-GEN <i>shakhṣ</i> <i>sb</i>	92

Table 2.8 Adjectival idioms

No.	Idiom	F.
1	ba‘īd-u l-ghawf-i far-NOM DEF-bottom-GEN lit. far-bottomed e.g. (of a feeling) deeply felt, (of a change) profound, (of a thought or analysis) deep	40
2	ṭawīl-u l-nafas-i long-NOM DEF-breath-GEN lit. long-breathed (of a policy, plan, or effort) capable of continuing for a long time, (of a person) patient or capable of continuing in what they are doing (e.g. negotiation) for a long time, or (of a text or speech) long-winded	37
3	nazīf-u l-yad-i clean-NOM DEF-hand-GEN lit. clean-handed acting ethically	13
4	wāsi’-u l-’ufuq-i broad-NOM DEF-horizon-GEN lit. broad-horizoned (of a person) broad-minded or (of an analysis or intellectual work) comprehensive or inclusive	19

Table 2.9 Adverbial idioms

No.	Idiom	F.
1	janb-a-n 'ilā janb-i-n side-ACC-INDF by side-GEN-INDF side by side (i.e. 'together' in the non-physical sense) <i>Side by side</i> is frequently used to mean 'together' in an abstract sense, as in the following example from the BNC: Now, the point about merit has to be recognized side by side with opportunities	940*
2	wajh-a-n li-wajh-i-n face-ACC-INDF to-face-GEN-INDF face to face (i.e. in non-physical confrontation) <i>Face to face</i> is often used to refer to non-physical confrontation as the following example from the BNC shows: Sometimes one is brought face to face with facts which cannot be buried.	555*

2.6.1 Notes on idiom frequencies

As the tables above show, some idioms only occur a few times, while others occur more frequently in the corpus texts. This may show that the ubiquity of idioms in language use that was pointed to in the preceding chapter does not entail that every idiom is common in use. Rather, it only means that this type of expression is abundant in spoken and written communication, yet with, sometimes strikingly, different frequencies for individual examples.

In order to account fully for the dissimilarities in the frequency of idioms, one would need to conduct an in-depth analysis of the characteristics of the contexts in which they occur. Such an endeavor is out of the question here. However, it is possible to point to some factors that seem to account, even in part, for why particular idioms are frequent or infrequent in the corpus. The question of how valid these explanations are, however, can only be answered by more research on the topic.

The meanings of some individual idioms may provide an explanation for why they are relatively frequent in the corpus. For instance, the following idiom occurs 122 times:

- (7) dāra-t raḥā *shay'*
 spin.PST-F quern.F *sth*
 sth (e.g. a war) took place in a violent way

The idiomatic meaning typically relates to wars and conflicts. Since these are a main focus of attention in newspapers, this could explain partially the high frequency of the idiom. Similar accounts could be proposed for the relatively high incidence

3 The semantic structure of Arabic idioms

This chapter is devoted to the analysis and discussion of the semantic structure of Arabic idioms. It is divided into four major sections. The first section briefly discusses the notion of *motivation*, which features in many works on idioms. The second presents the results of the analysis with regard to the figurative patterns that underlie the data examples. The third introduces the notion of *isomorphism* and looks at how Arabic idioms may be classified with regard to it. Finally, the fourth section discusses the results of the analysis.

Examining the semantic properties of idioms may shed light on different aspects of their discursive behavior, e.g. the nature and range of the discursive functions they perform. For instance, Moon (1998) found that 89 percent of the expressions in the class designated *metaphors** in her data have some evaluative content. As a step further in this direction, examining the figurative images underlying evaluative idioms could help explain how and why they differ in respect of their orientation, to convey positive, negative, or ambivalent attitudes (see Moon 1998 and Fernando 1996).

Furthermore, previous research on idioms shows that examining their semantic structure is a crucial step in explaining their formal behavior as manifested particularly in their potential for lexical and syntactic variation. For instance, Riehemann (2001) found that certain types of syntactic variation, such as the modification of individual idiomatic constituents and passivization, are only possible with semantically decomposable idioms.

3.1 Motivation

The term *motivation* refers to the relationship between the literal meaning of the idiom and its non-literal reading. As Geeraerts (2003: 436) points out, this notion has two facets: global and constituential. The former refers to:

the relationship between the original meaning of the idiomatic expression as a whole and its derived meaning. [The latter . . .] involves the relationship between the original, literal meaning of the constituent parts of the idiomatic expression, and the interpretation that those parts receive within the derived reading of the expression as a whole [if they receive any].

As Langlotz (2006) puts it:

An idiom reflects *global motivation* if the semantic extension from the literal to the figurative scene is still transparent. Thus, for *rock the boat* the relationship between the literal scene (ROCK THE BOAT) and its figurative interpretation (SPOIL A COMFORTABLE SITUATION) seems well-motivated: when someone rocks a boat, the boat is likely to be overturned; this, obviously, spoils the more satisfactory previous state or situation.

(p. 111, his italics and capitals)

Constituent motivation appears with idioms whose constituents possess *lexicalised figurative senses* that also appear outside the phrasal context of the idiom. [. . .]. For instance, *swallow* has the lexicalised figurative sense “accept patiently”. This meaning is present in idioms like *swallow the bitter pill* (accept an unpleasant fact) or *strain at a gnat and swallow a camel* (accept a major wrong thing while being concerned with a minor one).

(p. 111, his italics)

He also notes that *the bitter pill* is used in the meaning “unpleasant fact” outside the context of the idiom (p. 118).

On the other hand, an example of an unmotivated idiom in English is *shoot the breeze*, where many native speakers cannot establish links between its literal and idiomatic readings.

A number of points are in order. First, some idioms do not have (full) literal meanings for many native speakers. This could take place for one or more of the following reasons. First, the idiom could contain one or more unique components that do not occur outside it and do not bear any meaning for most native speakers today. These items are often termed *cranberry elements* (from *cran-* in *cranberry*, which is opaque to most native speakers of English). Second, the idiom could be syntactically idiosyncratic. That is, it could belong to the class that Fillmore *et al.* (1988) name *extragrammatical idioms* wherein the structures of the expressions “are not made intelligible by knowledge of the familiar rules of the grammar and how those rules are most generally applied” (p. 505). Finally, the idiom could contain a highly specialized lexical item that, though it occurs outside the context of the idiom in certain domains, is not known to most general language users. Idioms showing one or more of these features are sometimes labeled *literally non-compositional idioms* (Langlotz 2006).

Second, even when the idiom has a synchronic, literal meaning, whether it is motivated or not for a certain speaker depends on the depth and width of their knowledge of the world. As a result, an idiom could be motivated for one speaker but not for another. Thus, a degree of subjectivity is unavoidable in this regard. For example, a speaker who has sufficient knowledge of the boxing scenario may appreciate the relationship between the literal and idiomatic meanings of *throw in the towel*. On the other hand, another speaker who does not have such knowledge may not be able to recognize this relationship.

Third, native speakers may assign the idiom a literal interpretation that is different from its etymological one. This “newly” assigned reading could have a different relationship with the idiomatic meaning, or it could fully or partially block the semantic motivation of the idiom, rendering the expression opaque. This phenomenon may occur, for example, because one (or more) of the idiomatic components has undergone a process of semantic change, is a homonym of another lexical item, or is polysemous.

Idioms containing such elements are sometimes called *idioms with garden-path constituents* (Langlotz 2006), as the presence of these items often “leads contemporary native speakers up the garden path” about the literal meaning of the idiom. A famous English example of this class is *kick the bucket*. According to Gibbs (1994):

[p]eople generally think that *kick the bucket* comes from the situation of a condemned man with a rope around his neck who suddenly has the bucket he is standing on kicked out from beneath him. However, the phrase *kick the bucket* probably comes from a method of slaughtering hogs where the animal was strung up on a heavy wooden framework and its throat cut. *Bucket* is thought to be an English corruption of *buquet*, a French word for the wooden framework that the hog kicked in its death struggles.

(pp. 276–277, his italics)

Fourth, given the characterization above, motivation does not imply that the idiomatic meaning of an expression can be predicted on the basis of its literal reading (Langlotz 2006). Finally, motivation can be lost or weakened. As Geeraerts (2003: 442) notes, this:

often results from cultural changes. More often than not, the background image that motivates the figurative shift is an aspect of the material or the immaterial culture of a language community – and when the culture changes, the imagistic motivation may lose its force.

3.2 Patterns of semantic extension in idioms

Idioms have often been regarded as examples of conventionalized uses of figurative language (Gibbs 1994). This standpoint suggests two phases in the life history of an idiom. First, the expression starts life as an instance of ingenious exploitation of one or more figures of speech such as metaphor and metonymy. Second, gradually, because of its repeated occurrence, the originally creative expression shades into an institutionalized unit and becomes part of the phraseological repertoire of the language.

The reverse could occur too. That is, a literal expression may first become institutionalized through repeated use, then undergo a process of semantic extension whereby it acquires a new non-literal meaning. In order for the expression in this case to be regarded as an idiom, the new form-meaning relationship would need

Table 3.1 Distribution of idioms over figurative patterns

	<i>V-SBJ idioms</i>	<i>VP idioms</i>	<i>NP idioms</i>	<i>PP idioms</i>	<i>AP idioms</i>	<i>ADVP idioms</i>	<i>Total</i>
Metaphor	2	26	9	4	8	2	51
Metonymy	2		1				3
Metaphor– metonymy interaction	3	14	3	5	1		26
Conventional knowledge		1	3				4
Hyperbole				2			2
Emblems			1				1

to become conventionalized in the speech community (for a relevant discussion, see Cowie 1983).

I have identified 93 idiomatic meanings for the 70 examples in the sample. The figures provided in Table 3.1 and the analyses to follow reflect the distribution of these meanings over different patterns of semantic extension rather than the distribution of the expressions themselves. It should be noted too that when an idiom has more than one meaning, it could be associated with more than one of the categories below when its meanings are based on different patterns of semantic shift.

Several mechanisms of semantic extension participate in the creation of the idioms in the data. These include metaphor, metonymy, the interaction of metaphor and metonymy, semantic extension that is based on conventional knowledge, hyperbole, and emblemizing.

In addition to the interaction of metaphor and metonymy which underlies many examples in the data, some idioms seem to be based on other types of interaction between patterns of semantic extension. For simplicity, I have not set a distinct class for each of these types of interaction. Rather, I have subsumed such examples under the patterns of figuration that appear to play the major part in their production. The involvement of any other pattern of semantic extension has been discussed as appropriate, e.g. see the idiom in (20).

In five cases, it has not been straightforward to characterize the basic figuration in the idioms using any single one of the mechanisms referred to above. Therefore, these examples have been gathered and discussed in a distinct section under *Special Cases*. These have not been incorporated in Table 3.1.

In determining the figurative patterns underlying the data, the study relies on scrutinizing the relationship obtaining between the idiom's literal meaning and the non-literal meaning(s) that it takes on in *Arabic corpus*. In this context, some points should be borne in mind. First, four idioms in the data are *literally non-compositional*, that is, they do not bear literal meanings for many native speakers of Arabic today. This has become evident particularly from my discussions with native Arabic speakers. These examples are:

- (1) qalaba zahr-a l-mijann-i li-*shakhṣ/shay'*
 turn.around.PST back-ACC DEF-shield-GEN to-*sb/sth*
 he became hostile to *sb/sth* after he was a friend or supporter of them/it
- (2) *shakhṣ/shay'*_i bi-rummat-i-*hi*_i
*sb/sth*_i with-worn.out.piece.of.rope-GEN-M.3SG.POSS_i
 all *sb*_i (e.g. a group of people), or the whole *sth*_i
- (3) allutayyā wa-llatī
 REL.F.SG.DIMIN and-REL.F.SG
 much ado, or lengthy difficult discussions
- (4) *shakhṣ/shay'*_i 'alā bakr-at-i 'ab-ī-*hi*_i
*sb/sth*_i on young.camel-F-GEN father-GEN-M.3SG.POSS_i
*sb/sth*_i (e.g. a group of people or things) all together

Examples (1), (2), and (3) contain cranberry elements that carry no meaning to many native speakers today. These items are *mijann*, *rummah*, and *allutayyā*, respectively. Moreover, the idiom in (3) is syntactically ill-formed, wherein two relative pronouns are coordinated with no relative clauses present. The idiom in (4) contains *bakrah*, which only occurs within specialized contexts and bears no meaning for many Arabic native speakers. Patently, the presence of these lexical elements and grammatical features renders the idioms unmotivated for native speakers.

Nevertheless, it should be mentioned that native speakers may be able to recognize a partial, literal interpretation in some of these cases. For instance, in (1) they may construct a meaning such as “to turn around the back of X towards *sb/sth*”, whereas the cranberry element represented by X remains inaccessible to them.

It might be argued that it is irrelevant from a synchronic point of view to establish the semantic relationships between the idiomatic meanings of such examples and the literal meanings they probably had for native speakers at some point in the past. However, establishing these relationships, whenever possible, has proved useful when attempting to explain certain aspects of the linguistic behavior of idioms today. Therefore, these examples have not been excluded.

Second, two idioms in the data contain garden-path constituents:

- (5) nār-u-n 'alā 'alam-i-n
 fire-NOM-INDF on mountain-GEN-INDF
 very well-known entity
- (6) ḥalqat-u-n mufragh-at-u-n
 ring-NOM-INDF cast.in.a.mold.PPTCP-F-NOM-INDF
 vicious circle

The idiom in (5) is possibly based on the fact that a fire that is set on a mountain, particularly during night, is very visible to people passing by, hence the semantic extension to the state of being very well known (see also section 4.1.7). However, the word *'alam* does not have the meaning “mountain” for many native speakers today. Rather, it is frequently used to refer e.g. to “flag” or “sign”. Therefore, even

well-educated native speakers may not be able to construct the full original, literal meaning and, as a result, may not be able to comprehend the relationship depicted above. In more technical parlance, the etymological motivation of the idiom may not be clear to them.

In (6), etymologically the literal meaning of the idiom is probably “a metal ring that is cast in a mold” and, therefore, is seamless. Classically, the expression was used, typically as a simile, in situations wherein it is impossible to decide on the favorite member of a group of people. In such situations, the group, e.g. the sons of a man, is considered similar to a seamless ring of which one cannot come to a conclusion on where it begins or ends. This meaning of the word *mufraġhah* is still in use. However, this word also denotes another related meaning which is “hollow”, and native speakers seem to think only of this latter meaning rather than of the former one when they are asked to re-motivate the expression. This, as a result, veils the etymological motivation of the idiom.

Third, the following idiom can evoke more than one scene, because the nominal component is polysemous:

- (7) yanfukh-u fī qirbat-i-n maqṭū‘at-i-n
 blow-IND into bagpipe-GEN-INDF cut.PPTCP-GEN-INDF
 he is trying in vain to convince other people to take certain actions to
 improve a situation

The noun *qirbah* can mean a “skin” or “bagpipe”. It seems that the former meaning is the one that often comes to the minds of native speakers when they are asked about the relationship between the literal and non-literal readings of the idiom. In this meaning, two contexts seem to be possible for the literal scene. First, as some of my informants point out, skins are often used as butter churns and inflating the skin is a necessary step before shaking the milk. If the skin is defective, the effort will be fruitless. The other, but rather remote, possibility is the use of inflated skins as swimming aids. This seems to be an old practice. Also, some classical Arabic proverbs are based on this latter context.

Playing music using a bagpipe is a different possible context for the literal reading. Although, as noted above, native speakers seem to think more often of the meaning “skin”, this music-related context is supported as a possible etymological context by two factors. First, some in-text manipulations of the idiom suggest that its literal meaning is related to playing bagpipe music. Second, since attempts to convince people normally take place by talking to them, an analogy could be drawn between failing to produce anything that can be heard by using a defective bagpipe and failing to get one’s opinions recognized or heard responsively. Therefore, I have systematically used “bagpipe” as a translation of *qirbah* throughout the book.

Fourth, one important point that follows from this discussion is that native speakers may be able to motivate the idiom, i.e. to link between its literal and non-literal meanings, without knowing (for certain) the etymological context in which the literal scene occurs or used to occur (see Langlotz 2006). This probably applies more clearly to:

(8)	ʾamsaka	l-ʿaṣā	min-a	l-munṭaṣaf-i
	hold.PST	DEF-stick	at	DEF-middle-GEN
	he struck a balance			

Fifth, it might be argued that the discussion of the figurative patterns which underlie Arabic idioms that are loan translations from e.g. English may not be of much relevance to the Arabic context. In some cases, the introduction of these expressions into Arabic might be regarded as a result of the laziness or lack of knowledge of the translator who did not try to find, or did not know about the existence of, an originally Arabic expression that can adequately convey the meaning of the English idiom.

However, a good reason may be provided for the discussion of the figurative patterns of such idioms. That is, native Arabic speakers may re-motivate these idioms on the basis of their experience of the world. This motivation may or may not match the motivation on which the idiom is based in English. It may be relevant here to refer to Geeraerts' (2003) remark that speakers tend to find semantic links between the literal and non-literal meanings of even the most opaque idioms (see his notion of *semantic back-formation*).

It is acknowledged that in some cases the remoteness of the origin of the idiom from the Arab culture may prevent the loan translation from becoming an established idiom in MSA. In this respect, Holes (2004) points out that some literal translations of e.g. English idioms "may simply be too opaque and remote from the experience of the average reader unfamiliar with the western source languages to have any chance of survival" (p. 315).

Indeed, the fact that native Arabic speakers abandon loan translations that are so remote from their experience may indicate that they only adopt loan translations that they can motivate in some way or another. This provides more support for the decision to examine the figurative patterns of such expressions here.

Finally, as Radden (2003) rightly points out, people's experiences and knowledge systems are subjective. Therefore, their "characterisations of semantic structures including figurative language may be different" (p. 408). A major area in which this subjectivity may become noticeable is deciding whether a particular figurative expression is metaphorical or metonymic. In many cases, the borders between metonymy and metaphor are not sharply defined, and it is a "commonly acknowledged fact that a very large number of extended meanings are not amenable to characterisation as either exclusively metaphors or exclusively metonymies with respect to the root meaning" (Riemer 2003: 381). For this reason, some native Arabic speakers might disagree on the details of some of the analyses provided below.

3.2.1 Metaphor

Metaphor is a pattern of semantic extension that is typically based on a relationship of *similarity* between two concepts that belong to two different domains of human experience. As Kövecses (2002) points out, in addition to real similarity, metaphors may be founded on "perceived resemblance and correlations in experience" (p. 146).

In the following, I provide examples of metaphorical idioms along with comments on their figuration. Then I give some other examples without further analysis:

- (9) 'afala najm-u *shakhṣ/shay'*
 set.PST star-NOM *sb/sth*
 the glory or fame of *sb/sth* ended
- (10) dāra-t raḥā *shay'*
 spin.PST-F quern.F *sth*
sth (e.g. a war or elections) took place in a violent way
- (11) bala'a l-ṭu'm-a
 swallow.PST DEF-bait-ACC
 he swallowed the bait
- (12) yaḥruth-u fi l-baḥr-i
 plow.PRS-IND in DEF-sea-GEN
 he is plowing the sand
- (13) *shakhṣ/shay'* sābaqa l-rīḥ-a
sb/sth race.PST DEF-wind-ACC
sb/sth ran very fast, *sb* acted very quickly, or *sth* (e.g. the performance of a company) progressed very quickly

In (9), the idiom is based on a perceived similarity between the end of the fame or glory of somebody or something and the setting of a shining star. In (10), the violence of an activity such as war is regarded as a form of pressure that crushes the people involved in it as a quern grinds grain. In (11), reacting naïvely in accordance to a provocation is regarded as similar to the reaction of the fish when it swallows the fisherman's bait.

The idiom in (12) can be classified under what Langlotz (2006) calls *blended metaphorical idioms* (p. 131). These are based on integrating two distinct domains of human experience in the image underlying the idiom. This hybrid character of the figuration often leads the expression to have no possible literal meaning or one that is very unlikely to occur. The idiom in hand mixes what people know about plowing and its goals in agriculture with what they know about seas and their nature, creating an imaginative scene that is, at least, very unlikely to occur in the real world. This blend is then used to convey the idiomatic meaning.

Finally, in (13), the meaning "to run very fast" could be attributed to the category of blended metaphorical idioms, since the competitor in the imaginative racing scene is the wind which cannot normally be raced. As for the more abstract meanings of acting fast and making progress, they seem to be further metaphorical abstractions that are based on the first concrete meaning.

Other examples that belong to the category of metaphorical idioms include:

- (14) 'amsaka l-'aṣā min-a l-muntaṣaf-i
 hold.PST DEF-stick at DEF-middle-GEN
 he struck a balance

- (15) rimāl-u-n mutaharrrik-at-u-n
 sand.PL-NOM-INDF move.APTCP-F-NOM-INDF
 quicksand, or an unreliable basis for *sth* such a plan or project
- (16) 'ibrat-u-n fī kawmat-i qashsh-i-n
 needle-NOM-INDF in stack-GEN hay-GEN-INDF
 a needle in a haystack
- (17) qimmat-u jabal-i l-jalīd-i
 top-NOM mountain-GEN DEF-ice-GEN
 the tip of the iceberg
- (18) bayna nār-ayni
 between fire-DU.INDF.GEN
 between the devil and the deep blue sea

3.2.2 Metonymy

Metonymy is a type of semantic extension in which the concepts conveyed in the literal and non-literal readings of the expression belong to the same domain of human experience. In other words, the two concepts in this case stand in a relationship of *contiguity* rather than similarity. Various types of metonymic association exist, including the relationships between the parts and the whole, the instrument and the action, the container and the contained, the cause and the effect, among others (see e.g. Saeed 2003).

Examples of this type in the data include:

- (19) inqabaḍa ṣadr-u shakḥ
 contract.PST chest-NOM sb
sb felt depressed
- (20) tajammada l-dam-u fī 'urūq-i shakḥ
 freeze.PST DEF-blood-NOM in vein.PL-GEN sb
 the blood froze in *sb*'s veins
- (21) al-yad-u l-'ulyā
 DEF-hand.F-NOM DEF-upper.F
 the giver of charity

In both (19) and (20), the (assumed) bodily reactions stand for their emotional causes. In (19), the tension experienced in the muscles of the chest (or maybe in the heart) stands for its cause: depression. In (20), “the emotional experience [of fear] is felt to be associated with assumed or real changes in body temperature” (Kövecses 2002: 71) in which this temperature becomes lower than usual. However, the literal meaning violates truth conditions – after all, blood does not freeze inside people’s veins when they are frightened. This could be ascribed to the involvement of a hyperbolic element. Finally, in (21), the upper hand stands for the person who gives charity in a part-whole relationship. The hand is the body part normally used in the action of giving.

3.2.3 *The interaction of metaphor and metonymy*

Metaphor and metonymy may interact to provide the basis for idioms. The interaction between these two patterns in expressions that belong to the domain of linguistic action has been studied by Goossens (2003). However, some examples of the interplay of metaphor and metonymy in my data do not appear to fit in any of the categories that he puts forward in his account. Indeed, although Goossens points out that the domain of “linguistic action is sufficiently complex” and that his “data base exhibits enough diversity to allow us to come up with the main patterns”, he notes that his data base “is restricted and can therefore not be expected to provide an exhaustive account of the possible interaction patterns” (p. 350).

Accordingly, this section is broken into two parts. The first deals with the data examples that can be categorized according to Goossens (2003), and, therefore, I rely here heavily on his characterization of the phenomenon. The second part deals with the examples that do not fit in his categorization.

3.2.3.1 *Metaphonymy*

Goossens (2003) uses the label *metaphonymy* as a covering term for the interaction between metaphor and metonymy (the term *metaphonimic* is sometimes used as an adjective). His characterization of the phenomenon seems to account for 19 of the 26 cases that feature interaction between these two patterns in the data. Three of the subtypes described in his work can be identified in Arabic idioms: *metaphor from metonymy*, *metonymy within metaphor*, and *metaphor within metonymy*.

3.2.3.1.1 METAPHOR FROM METONYMY

In characterizing this type, Goossens (2003) notes that “the main point here is that underlying the metaphor there is an awareness that the donor domain and the target domain *can* be joined together naturally in one complex scene, in which case they produce a metonymy” (p. 366, his italics). In other words, the scene expressed in the literal interpretation of the expression and the situation or action conveyed in the idiomatic reading can co-occur in reality. This constitutes the metonymic basis in this type of expressions.

Among the examples Goossens (2003) gives to illustrate this pattern is the English idiom *beat one’s breast*, i.e. to “make a noisy open show of sorrow that may be partly pretence. Here the metonymic basis is the religious practice of beating one’s breast while one confesses one’s sins publicly” (p. 362).

However, he points out that in some cases the semantic extension in one and the same expression may be considered purely metaphorical or as involving a metaphor from metonymy. Indeed, “metaphor from metonymy occurs with varying degrees of cognitive saliency” (p. 358) and its “boundary lines with pure metaphors and pure metonyms are sometimes a little hazy” (p. 361, and see Radden’s remark above).

Sixteen idiomatic meanings can be classified under the metaphor from metonymy category. These belong to nine verbal idioms and three prepositional

ones. In what follows, I provide examples of these along with comments on their figuration. Then, I list some other examples of this type without further analysis:

- | | | | |
|------|---------------------------|------------------|---------------|
| (22) | 'adāra | l-khadd-a | l-'ākhar-a |
| | turn.PST | DEF-cheek-ACC | DEF-other-ACC |
| | he turned the other cheek | | |
| (23) | 'arghā | wa-'azbada | |
| | foam.PST | and-froth.PST | |
| | he fumed with rage | | |
| (24) | min | warā'-i | zahr-i |
| | from | behind | back-GEN |
| | behind <i>sb</i> 's back | | |
| (25) | taḥta | l-'aḏwā'-i | |
| | under | DEF-light.PL-GEN | |
| | in the spotlight | | |

In (22), the actual turning of the other cheek after being slapped on the first one is connected, at least given its Biblical reference, to refraining from retaliating (and forgiveness) after an insulting bodily attack. When the expression is used to denote refraining from retaliating in general without intending its literal meaning, it turns into a metaphor.

In (23), some people produce a lot of saliva that might come out of their mouths while shouting because of anger. This may be the metonymic basis of this idiom, wherein the effect stands for its cause. When it is used without intending its literal meaning, it becomes metaphorical and conveys only the meaning "act in a very angry way". Another possible account is that it is based on the image of an animal that, while angry, produces large amounts of saliva which come out of its mouth. This is perhaps the basis of the English idiom *froth at the mouth*.

In (24), when trying to keep somebody unaware of an activity, people may engage in that activity literally behind their back, hence the metonymic association between the actual involvement in an activity behind someone's back and keeping them in the dark. This metonymic linkage may be regarded as one in which the cause stands for the effect. When the expression is used without intending its literal meaning, it becomes metaphorical.

In (25), when someone is literally in the spotlight e.g. in a party or festival, this draws people's attention to them. This may provide a metonymic basis for the idiom wherein the cause stands for the effect. The expression becomes metaphorical when it is used to refer to being the center of public attention without intending its literal meaning.

Other examples that belong to this category include:

- | | | | | |
|------|--------------------------------------|----|---------|---------------------|
| (26) | waqafa | fī | ṭarīq-i | <i>shakhṣ/shay'</i> |
| | stand.PST | in | way-GEN | <i>sb/sth</i> |
| | he stood in the way of <i>sb/sth</i> | | | |

- (27) qalaba zahr-a l-mijann-i li-*shakhṣ/shay'*
 turn.around.PST back-ACC DEF-shield-GEN to-*sb/sth*
 he became hostile to *sb/sth* after he was a friend of or supporter to them/it
- (28) saraha wa-mariha
 move.freely.PST and-enjoy.oneself.PST
 he did as he liked
- (29) 'ala l-ṭāwilat-i
 on DEF-table-GEN
 on the table

A corollary of the foregoing characterization is that, in principle, idioms that are based on this type of figuration can be used as pure metonymies. That is, they may be used to refer to a scene where both the literal and idiomatic readings are relevant. As Goossens (2003) puts it, “these items have a hybrid character, in that they are metonyms in some contexts, metaphors from metonymy in others and sometimes undecided between these two interpretations in actual contexts” (p. 358).

However, in actual use, not all these expressions occur as metonymies. Examples that are used sometimes as metonymies include the idioms in (23), (24), and (26). With regard to (23), for instance, it can be interpreted as a metonymy when it is used with a human subject. However, in some contexts, even when the subject is human, it is not clear whether the literal meaning is intended or not, that is, in such cases both the metonymic and metaphor from metonymy interpretations are possible. Also, a prominent example in this context is (24). In the following corpus extract, both its literal and idiomatic meanings can be relevant in the football match scene:

- (30) istikhḏām-a l-khushūnat-i wa-l-ḏarb-i bidūn-i
 using-ACC DEF-roughness-GEN and-DEF-beating-GEN without-GEN
 kurat-i-n min khalf-i zahr-i l-ḥakam-i (Al-Ahram, 1999)
 ball-GEN-INDF from behind-GEN back-GEN DEF-referee-GEN
 lit. violent conduct and off-the-ball fouls behind the referee's back

3.2.3.1.2 METONYMY WITHIN METAPHOR

According to Goossens (2003), in this type, “a metonymically used entity is embedded in a (complex) metaphorical expression. The metonymy functions within the target domain” (p. 367). There are both metaphorical and metonymic relationships between the literal and idiomatic readings. The metonymic relationship is present within a metaphor that functions at the level of the meaning of the whole expression. The metonymic item is actually a shared element in both the source and target domains of the metaphor. It “functions metonymically in the target domain only, whereas it is interpreted literally or (more often) (re)interpreted metaphorically in the donor domain” (Goossens 2003: 363).

One of Goossens' examples of this pattern is *shoot one's mouth off*, which he defines as: “talk foolishly about what one does not know about or should not talk about” (p. 364). He analyzes the figuration as follows:

The donor domain is the foolish or uncontrolled use of firearms: the foolish (and therefore potentially, though not intentionally, dangerous) use of a gun is mapped onto unthoughtful linguistic action. By integrating *mouth* into a scene relating to the use of firearms it is reinterpreted as having properties of a gun in the donor domain; this is the metaphorisation in the donor domain. In the recipient domain, however, there is a first level of interpretation which amounts to something like “using one’s mouth foolishly,” in which *mouth* is a metonymy for *speech faculty*. [. . . The] significance of the metonymy becomes clear, if one replaces *mouth* by parts of the body which are less or not functional in the act of speaking (such as *nose* or *eyes*).

(p. 364, his italics)

Two idioms may be regarded as metonymies within metaphors:

- | | | | |
|------|------------------------------|-----------------------|-----------------------|
| (31) | <i>lā</i> | yuṣaddiq-u | ‘ayn-ay- <i>hi</i> |
| | NEG | believe.PRS.M.3SG-IND | eye-DU.ACC-M.3SG.POSS |
| | he does not believe his eyes | | |
| (32) | nazīf-u | l-yad-i | |
| | clean-NOM | DEF-hand-GEN | |
| | acting ethically | | |

In (31), at the highest level, i.e. the level of the embedding metaphor, this idiom features a metaphorical extension that is possibly based on a perceived similarity between the reaction of a person who is very surprised by something they see and that of a person who does not accept what someone tells them as true. The word ‘aynān “two eyes” then is interpreted metaphorically in the source domain by attributing a human quality to the eyes, i.e. the quality that a human being may or may not be believed as a source of information. In the target domain, i.e. the domain of the idiomatic meaning, on the other hand, the eyes, as the organ of vision, stand in a metonymic relationship with the what-is-seen part of the idiomatic reading.

The idiom in (32) is based on a combination of a metaphor in which what is ethical is considered as clean and a metonymy in which *yad* “hand” stands for the activity (see Kövecses 2002: 210, for an analysis of *have clean hands*).

3.2.3.1.3 METAPHOR WITHIN METONYMY

In this case, an element that is used metaphorically is embedded into a metonymy. Goossens (2003) illustrates this type with the English idiom *be/get on one’s hind legs*, which he defines as: “stand up in order to say or argue something, esp. in public” (p. 366). His analysis of this example goes as follows:

The peculiarity about this item is perhaps best revealed if we leave out *hind: being/getting up on one’s legs* with reference to “standing up in order to say something in public” is metonymic, there is an overall sense of somebody

standing up *and* saying something publicly. The addition of *hind* forces us to reinterpret the expression in terms of an animal standing up. This suggests a greater effort, an event which attracts more attention. At the same time there is a bathetic effect, because a human being is interpreted as being involved in the pseudo-achievement of standing on two legs. One may, of course, also argue that the addition of *hind* makes the expression as a whole metaphorical; it is only to the extent that we process it with an awareness of the metonymy, that it is more adequate to view this as a metaphor embedded into a metonymy. (p. 366, his italics)

Only one idiom in the data could be classified as a metaphor within metonymy:

- (33) al-yad-u l-‘ulyā
 DEF-hand.F-NOM DEF-upper.F
 the upper hand

yad “hand” stands metonymically for power, as it is the instrument with which power is often exercised. The metaphorical extension of ‘*ulyā* “upper” is based on perceiving being in control in terms of being at a higher physical position (for an analysis of the figuration in the English idiom *gain the upper hand* see Kövecses 2002: 210).

3.2.3.2 *Another type of interaction of metaphor and metonymy*

Seven idiomatic meanings reflect an interaction between metaphor and metonymy that does not straightforwardly fit the foregoing classification. Although these cases have been discussed separately here, since the term *metaphonymy* refers to the interaction of metaphor and metonymy in general, these examples can, of course, be subsumed under it. Indeed, Langlotz (2006) describes some similar expressions in English as metaphonymic.

These meanings belong to three verb-subject idioms, two nominal idioms, and one prepositional idiom:

- (34) *lā* taqūm-u li-*shakhṣ/shay*’ qā’imat-u-n
 NEG stand.up.PRS.F-IND for-*sb/sth* pillar.F-NOM-INDF
sb/sth is completely ruined (e.g. socially or professionally)
- (35) *shay*’_i ‘akhadha ba‘d-u-*hu*_i bi-riqāb-i ba‘d
*sth*_i hold.PST some.of-NOM-POSS_i at-neck.PL-GEN some
*sth*_i (e.g. a set of ideas or problems) interconnected
- (36) tajammada l-dam-u fī ‘urūq-i *shay*’
 solidify.PST DEF-blood-NOM in vein.PL-GEN *sth*
sth (e.g. an organization or political movement) became inactive
- (37) shadd-u-n wa-jadhb-u-n
 tugging-NOM-INDF and-pulling-NOM-INDF
 great tension or struggle (e.g. in a strained political or social state)

- (38) nār-u-n ‘alā ‘alam-i-n
 fire-NOM-INDF on mountain-GEN-INDF
 very well-known entity
- (39) ‘alā ṣafīḥ-i-n sākhin-i-n
 on sheet.iron-GEN-INDF hot-GEN-INDF
 (of a person) very anxious, or (of a situation) very tense or instable

Evidently, these are metaphors, as the literal and idiomatic meanings in each case belong to two different domains, i.e. they are not contiguous, and the meaning extensions involve a relationship of similarity. However, they are not pure metaphors, since there seems to be a set of metonymic shifts that took place prior to the metaphorical extensions.

At first glance, the figuration in these idioms might be attributed to the metaphor from metonymy type discussed above. Yet, a closer look helps see the dissimilarity between the two cases. According to Goossens (2003), the main point in the metaphor from metonymy type is that “the donor domain and the target domain *can* be joined together naturally in one complex scene, in which case they produce a metonymy” (p. 366, his italics). This is not the case in these idioms. They cannot be used as pure metonymies wherein both their literal and idiomatic readings are relevant. For instance, in (34), being socially or professionally ruined is not contiguous with “a tent or building having none of its pillars up”.[†] In (37), “social or political tension or struggle” does not naturally co-occur with concrete pulling and tugging. And, in (38), fame is not contiguous with “a fire on a mountain”.

In these cases, the perceived similarity relationships obtain between the idiomatic reading and an implication of the literal meaning rather than the literal meaning itself. These relationships are basically between social ruination and physical ruination in (34), ideational or abstract interconnection and physical interconnection in (35), inactivity and death in (36), social tension and physical tension in (37), well-knownness and clear visibility in (38), and emotional or social instability and physical instability in (39). Many of these cases, and several other idioms in the data, relate to more general conceptual metaphors. For example, the idiom in (38) may be connected to a general model in which “knowing” is perceived in terms of “seeing”.

3.2.4 Conventional knowledge

Shared knowledge about the world within a given speech community may provide the basis for some idioms (Kövecses 2002). The sources of this shared knowledge in the data vary. They include general knowledge of the world, religion, and, even, superstitions. The three examples that can be classified under this category are:

- (40) ḥibr-u-n ‘alā waraq-i-n
 ink-NOM-INDF on paper-GEN-INDF
 inactive (e.g. of decisions, laws, or agreements)

From what we know about the world, agreements, truces, etc. are usually written, so when these are “only ink on paper”, this implies that they are not in (full) force.

- | | | | |
|------|----------------|----------|-------------------------|
| (41) | 'akmala | nişf-a | dīn-i-hi |
| | perfect.PST | half-ACC | religion-GEN-M.3SG.POSS |
| | he got married | | |

The idiomatic meaning in this case is derived from knowledge of some Islamic teachings. In the third case, the idiomatic meaning “Adam’s apple” is based on a superstition, and the meaning “the forbidden fruit” is based on a particular interpretation of some religious texts:

- | | | |
|------|--|-------|
| (42) | tuffāḥat-u | 'ādam |
| | apple-NOM | Adam |
| | Adam’s apple, or (esp. in literary language) the forbidden fruit | |

With regard to the first meaning, the Oxford English Dictionary (OED) notes that this part of the throat is “so called from a superstitious notion that a piece of the forbidden fruit stuck in Adam’s throat, and occasioned this prominence.”

3.2.5 *Hyperbole*

The term *hyperbole* is used for “exaggeration or overstatement, usually deliberate and not meant to be taken (too) literally” (McArthur 1998: “hyperbole”). The following idioms can be considered instances of this trope:

- | | | | |
|------|---|-----------------------------|---------------------|
| (43) | fī | 'a'qāb-i | <i>shakhṣ/shay'</i> |
| | in | heel.PL-GEN | <i>sb/sth</i> |
| | immediately behind <i>sb</i> , or immediately after <i>sb/sth</i> | | |
| (44) | 'alā | marmā | ḥajar-i-n |
| | at | distance.that.sth.is.thrown | stone-GEN-INDF |
| | of proximity in space or time | | |

In (43), the spatial meaning is probably based on the presence of a deliberate exaggeration. The temporal reading is a further metaphorical extension that is based on the general model in which temporal relationships are perceived in terms of spatial ones (see Lakoff and Johnson 2003 and Kövecses 2002).

As to the spatial meaning in (44), when the entity or location is actually within a stone’s throw, the expression would be literal. Nevertheless, it is doubtful that native speakers do mean the literal reading even in this case. When the entity or location is close but not really within a stone’s throw, the expression could be regarded as an instance of hyperbole. As for the temporal meaning of the idiom, this is driven by a further metaphorical extension which is based on its spatial meaning.

Following Burger (1998), Langlotz (2006) terms semantic shifts that are based on other semantic extensions *secondary metaphorization*, pointing out that

“[i]diomatic polysemy reflects the creation of an institutionalised pattern of meaning extension based on a figurative construction. [...] Such lexicalised meaning variants therefore reflect the general potential for meaning shift in idioms” (p. 181).

3.2.6 Emblems

The OED defines *emblem* as a “picture of an object (or the object itself) serving as a symbolical representation of an abstract quality, an action, state of things, class of persons, etc.” The emblematic relationship underlies the meaning extension in:

- (45) ghuṣn-u l-zaytūn-i
 branch-NOM DEF-olive-GEN
 olive branch

3.2.7 Special cases

Five idioms show characteristics that make it difficult to categorize them under any single one of the classes above:

- (46) dam-u-n ’azraq-u
 blood-NOM-INDF blue-NOM
 blue blood
- (47) *shakhṣ/shay*’_i ‘alā bakr-at-i ’ab-ī-*hi*_i
shakhṣ/shay’_i on young.camel-F-GEN father-GEN-M.3SG.POSS_i
*sb/sth*_i (e.g. a group of people or things) all together
- (48) *shakhṣ/shay*’_i bi-rummat-i-*hi*_i
*sb/sth*_i with-worn.out.piece.of.rope-GEN-M.3SG.POSS_i
 all *sb*_i (e.g. a group of people), or the whole *sth*_i
- (49) allutayyā wa-llatī
 REL.F.SG.DIMIN and-REL.F.SG
 much ado, or lengthy difficult discussions
- (50) marbaṭ-u l-faras-i
 place.where.sth.is.tied.up-NOM DEF-mare-GEN
 the central point (e.g. in a problem or argument)

As to the idiom in (46), the following extracts from the OED about *blue blood* are enlightening:

Blood is popularly treated as the typical part of the body which children inherit from their parents and ancestors; hence that of parents and children, and of the members of a family or race, is spoken of as identical, and as being distinct from that of other families or races.

Blue blood: that which flows in the veins of old and aristocratic families, a transl. [i.e. translation] of the Spanish *sangre azul* attributed to some of the

oldest and proudest families of Castile, who claimed never to have been contaminated by Moorish, Jewish, or other foreign admixture; the expression probably originated in the blueness of the veins of people of fair complexion as compared with those of dark skin; also, a person with blue blood; an aristocrat.

Therefore, *blood* may be linked metonymically to the race or family. *Blue* relates to the group of aristocrats to whom the expression historically referred, i.e. because of the blueness of their veins. Both *blood* and *veins* are connected, since the latter is the container of the former. The literal reading of the idiom, however, violates truth conditions, as blood cannot be blue. Finally, it seems that at some point, probably due to some similarity in status and/or behavior, the narrow meaning of the expression as used within its etymological context was extended through a generalization process to refer to any race or group of people that is considered noble or aristocratic. Because of this generalization, the literal meaning of *blue* does not relate to the current idiomatic reading, as they are not contiguous with each other.

However, the Arabic idiom seems to be a loan translation of the English, or maybe other Western, version. Therefore, the foregoing analysis might not be relevant in its entirety to the Arabic context. The reason for this is that it is possible, at least in principle, that the expression was borrowed after the probable generalization process referred to had taken place (see also section 4.1.7).

In (47), this classical literally non-compositional idiom possibly originates in the story of three sons who were killed by an enemy during a journey in the desert, and then their bodies were sent back home carried “all together” on their father’s camel. The idiom therefore has a metonymic basis. It seems that at a later stage, a generalization process took place, whereby the expression has become to mean “all together” in a wide range of situations, without any reference to its original context.

In (48), this literally non-compositional idiom belongs to the context of selling and buying farm animals. The word *rummah* means “a worn out piece of rope” that could be used as an animal’s lead. Thus, a sentence such as:

- | | |
|----------------------|---------------------------------|
| (51) khudh-hu | bi-rummat-i-hi |
| take.IMP.M.3SG-M.3SG | together.with-a.worn.out.piece. |
| | of.rope-GEN-M.3SG.POSS |

would literally mean “take it together with its worn out lead”. This shows that the idiom has a metonymic basis. A generalization process seems to have taken place, by which the expression has become to mean “all” or “completely” in a range of contexts that is much wider than its original one.

Both (47) and (48) may be regarded as cases of grammaticalization, i.e. the change whereby lexical items and constructions come in certain linguistic contexts to perform grammatical functions (Hopper and Traugott 2003). In the present cases, the expressions have abandoned their original meanings to denote a

meaning of quantification. On the role of metonymy in the occurrence of such phenomena, Brinton and Traugott (2005) point out that:

[w]hile the result of grammaticalization is often synchronically metaphorical, textual evidence for the development of many grammatical formatives out of lexical and constructional material is metonymic in the sense that it is highly context-bound and arises out of the implicatures in the speaker-hearer communicative situation (p. 28).

They note that “implicatures from lexical items in specific, repeatedly used contexts come to be semanticized” (p. 29). Also, Bennett (2002) points out that “[o]ne mechanism for semantic change is for an inference to be reinterpreted as part of the meaning of an expression” (p. 10). In (47) and (48), it seems that the inferences had even been reinterpreted as the whole meanings of the expressions.

The study has arrived at no satisfactory conclusions regarding the nature of the semantic extension underlying the idioms in (49) and (50). However, as far as the former is concerned, one plausible account, which has been referred to in some classical Arabic dictionaries, is that the expression is the remnants of two coordinated, complete relative clauses that denoted a very unpleasant state of affairs.

3.3 Arabic idioms and isomorphism

According to Geeraerts (2003), the term *isomorphism* refers to:

a one-to-one correspondence between the formal structure of the expression and the structure of its semantic interpretation, in the sense that there exists a systematic correlation between the parts of the semantic value of the expression as a whole and the constituent parts of that expression. [. . . Isomorphism is a static notion] that merely notes that a one-to-one correspondence between the parts of the semantic value of the expression as a whole and the meanings of the constituent parts of the expression can be detected, regardless of the question whether this correspondence has come about through a process of bottom-up derivation or through a top-down interpretative process (p. 437).

Examples of isomorphic idioms in English include *to kill two birds with one stone* – “to achieve two goals in one action” – and *to spill the beans* – “to reveal secret information”. The constituent parts of the idioms in these cases can be linked to parts of their meanings. In the former example, *kill* can be linked to “achieve”, *two birds* to “two goals”, and *one stone* to “one action”. In the latter, *spill* can be connected to the action of “revealing” and *the beans* to “the information” disclosed.

Some examples of non-isomorphic idioms in English are *saw logs* – “snore” – and *on edge* – “tense”. In both cases, although the relationship between the literal and idiomatic meanings is transparent, no connections can be made between the formal

constituents of the idioms and identifiable parts of their meanings. The first example consists of the two-place verb *saw* and *logs* which functions as its patient. The idiomatic meaning, on the other hand, is denoted by the one-place verb *snore*; therefore *logs* has no place in the semantic structure of the idiomatic meaning. The second example consists of the locative preposition *on* and the location *edge*. Since the idiomatic meaning is an attribute that is denoted by an adjective, i.e. *tense*, the formal structure of the expression does not map onto its semantic structure.

Isomorphism overlaps, as Langlotz (2006) points out, with the notion of *decomposition*, which is a directional top-down process whereby identifiable parts of the idiomatic meaning can be assigned to the constituent parts of the idiom. In this regard, Nunberg *et al.* (1994) classify idioms into two categories: *idiomatic combinations* and *idiomatic phrases* (these two types are called *decomposable* and *non-decomposable idioms* respectively in other works, e.g. Villavicencio and Copestake 2002, and for a justification for using these labels see Riehemann 2001). In the former, the non-literal meaning can be distributed over the idiomatic components by assigning a non-standard meaning to each of them. In the latter case, on the other hand, such distribution of the idiomatic meaning over the parts of the expression is not possible.

Using the English idioms *pull strings* as an example of the former type, Nunberg *et al.* (1994) note that one would:

be able to establish correspondences between the parts of the structured denotation of the expression (the relation of exploiting, the connections exploited) and the parts of the idiom (*pull* and *strings*), in such a way that each constituent will be seen to refer metaphorically to an element of the interpretation. That is, the idiom will be given a compositional, albeit idiosyncratic, analysis (p. 296).

However, this is not to say that the meanings of the idioms are completely equal to those of their literal paraphrases. The figurative nature of many idioms enables them to evoke mental images and implications that may not be attainable through any other literal means (see Croft and Cruse 2004). In the context of the conceptual view of idioms, Gibbs (1994) illustrates this point using the English idiom *spill the beans*:

phrases like *spill the beans* cannot simply be paraphrased as “to reveal a secret” in the way most idiom dictionaries give definitions [. . .]. The mapping of source domains, such as containers, onto target domains, such as minds, results in very specific entailments about the act of revealing a secret. Thus, the act of revealing a secret is usually seen as being caused by some internal pressure within the mind of the revealer, the action is thought to be done unintentionally, and the action is judged as being performed in a forceful manner [. . .]. One interesting possibility is that people actually draw on these inferences about the act of revealing a secret each time they comprehend the idiom *spill the beans*. If this is the case, people might be less likely to draw such inferences about causation, intentionality, and manner when

comprehending literal paraphrases of idioms. Literal paraphrases, such as *reveal the secret*, are not motivated by the same set of conceptual metaphors as are specific idioms, such as *spill the beans*. For this reason, people do not view the meanings of *spill the beans* and *reveal the secret* as equivalent, despite apparent similarity (p. 303).

In order to decide whether an idiom is isomorphic or not, a body of interdependent information needs to be at the analyst's disposal. This includes (1) a precise description of the meaning(s) that the expression takes on in different contexts, (2) information on the presence of any "discourse entities" (Pulman 1993) or "discourse referents", whether stated explicitly in the text or not, that can be linked to the idiomatic constituents in a one-to-one fashion, and, very importantly, (3) determining whether or not the idiomatic constituents are used in their figurative senses outside the context of the idiom, since, if they are, this would be a clear indication of the idiom's isomorphism (see e.g. Geeraerts 2003).

Analyzing the data obtained from *Arabic corpus*, and other sources, has been of much help in providing this information. The analysis shows that it is not sufficient to classify Arabic idioms as either isomorphic or not. A third class is needed to subsume examples that are used in both ways. Tables 3.2 and 3.3 show the distribution of the data over these three classes. In the former, the classification is produced

Table 3.2 Isomorphism and the syntactic types of idioms

	<i>V-SBJ idioms</i>	<i>VP idioms</i>	<i>NP idioms</i>	<i>PP idioms</i>	<i>AP idioms</i>	<i>AdvP idioms</i>	<i>Total</i>
Isomorphic	2	20	10	2	9		43
Non-isomorphic	5	16	10	11		2	44
Isomorphic/ non-isomorphic		5	1				6

Table 3.3 Isomorphism and the patterns of semantic extension in idioms

	<i>Isomorphic</i>	<i>Non-isomorphic</i>	<i>Isomorphic/ non-isomorphic</i>
Metaphor	37	10	4
Metonymy		3	
Metaphor-metonymy interaction	5	19	2
Conventional knowledge		4	
Hyperbole		2	
Emblems		1	
Special cases	1	5	
Total	43	44	6

with reference to the syntactic types of the idioms, whereas in the latter, it is produced with regard to their patterns of semantic extension. The examples discussed under *Special Cases* above have been considered in both tables. In the latter, they have been incorporated in a separate column. Once more, the numbers provided below reflect a classification of the idiomatic meanings with regard to isomorphism rather than a classification of the expressions themselves (see section 3.2).

Examples of isomorphic idioms include:

- (52) 'afala najm-u shakhṣ/shay'
 set.PST star-NOM sb/sth
 the fame or glory of *sb/sth* ended
- (53) bala'a l-tu'm-a
 swallow.PST DEF-bait-ACC
 he swallowed the bait
- (54) ḥāraba ṭawāḥīn-a l-hawā'-i
 fight.PST mill.PL-ACC DEF-air-GEN
 he tilted at windmills
- (55) dassa l-summ-a fi l-'asal-i
 put.secretly.PST DEF-poison-ACC in DEF-honey-GEN
 he deceptively promoted something bad (e.g. a way of thinking or behaving)
 by presenting it in something else good
- (56) qimmat-u jabal-i l-jalīd-i
 top-NOM mountain-GEN DEF-ice-GEN
 the tip of the iceberg
- (57) dam-u-n 'azraq-u
 blood-NOM-INDF blue-NOM
 blue blood
- (58) rimāl-u-n mutaharrik-at-u-n
 sand.PL-NOM-INDF move.APTCP-F-NOM-INDF
 quicksand, or an unreliable basis for *sth* such as a plan or project
- (59) baqarat-u-n ḥalūb-u-n
 cow.F-NOM-INDF milch-NOM-INDF
 milch cow
- (60) 'ibrat-u-n fi kawmat-i qashsh-i-n
 needle-NOM-INDF in stack-GEN hay-GEN-INDF
 a needle in a haystack
- (61) bayna nār-ayn
 between fire-DU.GEN-INDF
 between the devil and the deep blue sea

For instance, in (52), *najm* “star” stands for the glory or fame of an entity, and the verb *'afala* “to set” may be interpreted as “to end”. In (53), *al-tu'm* “the bait” stands for the provocation, and the verb *bala'a* “to swallow” idiomatically conveys the meaning “to react naïvely”. In (55), *al-summ* “the poison” stands for the bad idea or pattern of behavior that the subject promotes, and *al-'asal* “the honey”

stands for the attractive package in which this idea is presented. In (59), *baqarah* “cow” stands for the source of profit, and *ḥalūb* “milch” for the quality of being productive and steady. Finally, in (61), *nār-ān* “two fires” refers metaphorically to the two difficult choices in the idiomatic meaning.

However, it should be noted that even when it is possible to link all the formal constituents to identifiable parts of the idiomatic meaning, some meaning parts may not have parallel formal constituents. For instance, in (60), an essential part of the meaning is that it is hard to locate the entity that the needle metaphorically represents. However, this essential part does not map onto any single constituent.

Examples of non-isomorphic idioms include:

- (62) *tajammada* *l-dam-u* *fī* *‘urūq-i* *shakhs*
 freeze.PST DEF-blood-NOM in vein.PL-GEN *sb*
 the blood froze in *sb*’s veins
- (63) *sābaqa* *l-rīḥ-a*
 race.PST DEF-wind-ACC
 he ran or acted very fast
- (64) *waqafa* *fī* *ṭarīq-i* *shay’*
 stand.PST in way-GEN *sth*
 he stood in the way of *sth*
- (65) *’arghā* *wa-’azbada*
 foam.PST and-froth.PST
 he fumed with anger
- (66) *saraḥa* *wa-mariḥa*
 move.freely.PST and-enjoy.oneself.PST
 he did as he liked
- (67) *nār-u-n* *‘alā* *‘alam-i-n*
 fire-NOM-INDF on mountain-GEN-INDF
 very well-known entity
- (68) *qarṣat-u* *’udhun-i-n*
 pinch-NOM ear-GEN-INDF
 a light non-physical punishment that serves as a warning of a possible cruel one
- (69) *ḥibr-u-n* *‘alā* *waraq-i-n*
 ink-NOM-INDF on paper-GEN-INDF
 inactive (e.g. of decisions or laws)
- (70) *fī* *’a’qāb-i* *shakhs/shay’*
 in heel.PL-GEN *sb/sth*
 close behind *sb*, or immediately after *sb/sth*
- (71) *min* *warā’-i* *zahr-i* *shakhs*
 from behind-GEN back-GEN *sb*
 behind *sb*’s back

Five idioms are used in both ways. For example, the idiom in (72a) can be used in the meaning “much social or political tension” wherein it is not isomorphic, as in

raff-i l-intizār-i (Al-Hayat, 1996)

shelf-GEN DEF-waiting-GEN

lit. it seems that the last governmental formation has excluded them and put them on the shelf of waiting

3.4 Discussion

This section discusses the following points as revealed in the analysis: (1) any correlations between the patterns of figuration and the syntactic types of idioms, (2) the high frequency of idioms that involve metaphorical extensions, (3) the high incidence of the metaphor from metonymy pattern compared with the other metaphonimic types, (4) why idioms that are based on the metaphor from metonymy pattern vary with regard to their use as pure metonymies, and (5) how the differences between idioms with regard to isomorphism can be explained.

The study shows that six out of the thirteen meanings of the prepositional idioms involve a metonymic element. These meanings belong to four idioms that involve an interaction between metaphor and metonymy and the two prepositional idioms that have been discussed among the five special cases. This correlation between prepositional idioms and the metonymic extension is not surprising in view of the range of the meanings that Arabic prepositions convey. They communicate various associations in time or space or express different relationships of causality, instrumentality, and accompaniment, *inter alia* (see e.g. Ryding 2005 and Badawi *et al.* 2004). Since metonymy is by definition concerned with meaning extensions that are based on such relationships, when prepositional phrases are used figuratively, this use often involves a metonymy.

The fact that idioms very often involve metaphorical extensions could be ascribed in part to the ubiquity of this pattern in thought and in language use in general (see Lakoff and Johnson 2003 and Gibbs 1994). Indeed, as noted earlier, many idioms in the data can be linked to some conceptual metaphors. For instance, the data contain examples that may be linked to general patterns of metaphorical mapping such as those between temporal relationships and spatial relationships, making progress and moving forward, social or emotional instability and physical instability, and knowing and seeing.

Additionally, metaphors fulfill several important communicative functions. In particular, they provide a compact way to convey complex meanings and enjoy a high degree of vividness (Gibbs 1994) that might not be found in literal language and in non-literal expressions that are based on other patterns of semantic extension. These communicative advantages could lead speakers and writers to prefer using particular metaphorical expressions over other possible alternatives, hence turning these expressions into idioms through institutionalization.

The metaphor from metonymy pattern is far more frequent in the data than the other metaphonimic types. In characterizing this pattern, Goossens (2003) notes that:

[m]etaphor from metonymy implies that a given figurative expression functions as a mapping between elements [A and B] in two discrete domains, but

that the perception of “similarity” is established on the basis of our awareness that A and B are often “contiguous” within the same domain. This frequent contiguity provides us with a “natural,” experiential, [*sic*] grounding for our mapping between two discrete domains (p. 368).

In fact, this experiential basis does not only provide the grounding for the mapping between the two distinct concepts, but it can also be regarded as a plausible explanation for the high frequency of this pattern as opposed to the other types, which do not share the same degree of embedding in human everyday experience.

As far as the metonymy within metaphor pattern is concerned, the condition to incorporate a shared element, i.e. one that functions both in the source and target domains, might at some level be considered a restriction on the productivity of this type, and hence provides at least a partial explanation for its infrequency in the data. However, in order to better understand this, other studies need to be conducted on different sets of data. In those studies, some effort should be made to categorize idioms in semantically coherent groups and to pay close attention to the question of how distinct the target domains are from the source domains. This would help uncover relevant patterns and provide a way to verify, or not, the generalizability of the findings of this study.

It is important in this context to note that Goossens (2003), in his study on figurative expressions conveying linguistic action, reported a high frequency of the metonymy within metaphor type. According to him, this was the case when the source domain involved a body part that “can be functional in linguistic action” (p. 365). This finding clearly shows the significance of the suggestion to assemble idioms in semantically consistent groups in preparation for a closer investigation.

Finally, the metaphor within metonymy type is represented by only one example in the sample. “The fact that if we embed a metaphor into a metonymy, it tends to ‘metaphorise’ the whole expression” (Goossens 2003: 367) may explain this rarity.

Not all the metaphor from metonymy idioms are used as pure metonymies. This may be due to linguistic and/or socio-cultural factors. For example, the idioms in (23) “he foamed and frothed”, (24) “behind *sb*’s back”, and (28) “he moved freely and enjoyed himself” can be used metonymically because they are clearly motivated and their metonymic basis is relevant to the present-day world of native speakers. On the other hand, an idiom such as (27) “he turned the back of the shield towards *sb*” is always used as a metaphor from metonymy. The reason for this is mainly linguistic, as the expression is literally non-compositional, thus it is not motivated for most native speakers. It might be added that even when this idiom is motivated for speakers who possess more linguistic knowledge, it is still used only as a metaphor from metonymy because, generally, the scene it evokes in its literal reading is not related to everyday life today.

In some other cases, restrictions on the purely metonymic uses might occur because of socio-cultural changes in the speech community. For instance, an expression that belongs to this type may continue to be used, whereas the situation with which it is metonymically associated has ceased to occur in the real world. Also, the metonymic basis of an idiom might be irrelevant to the everyday life of

The metonymy highlights an entailment in the literal-meaning domain. Thus, *go around in circles* entails stagnation. This entailment is then metaphorised to the more abstract idiomatic meaning: *go around in circles* denotes stagnation in any type of target activity. Since this process of figuration projects a metonymic association-pattern onto the target-domain, this idiom, as well as the other idioms of this class, does not reflect direct constituentual mapping (p. 129).

Nevertheless, some idioms which involve metonymies are isomorphic, e.g.:

- | | | | |
|------|--|----------------------|----------------------------|
| (75) | al-yad-u
DEF-hand-NOM
the upper hand | l-‘ulyā
DEF-upper | |
| (76) | dam-u-n
blood-NOM-INDF
blue blood | ’azraq-u
blue-NOM | |
| (77) | rafa‘a
raise.PST
he supported <i>sth</i> | rāyat-a
flag-ACC | <i>shay’</i>
<i>sth</i> |

The analyzability of the first two examples is due to the fact that their metonymic extensions are constituentual rather than global. The noun in (75) and the noun and adjective in (76) are motivated, at least as far as their etymology is concerned, by individual metonymic relationships and stand, each on its own, for other contiguous concepts. This constituentual motivation by separate patterns of figuration helps explain why these elements map onto identifiable parts of the idiomatic meaning, rendering the examples isomorphic.

The idiom in (77) has been classified as a metaphor from metonymy. People do wave the flags e.g. of their countries in the context of showing support for them. The metonymic extension thus is global. However, an emblematic relationship can be noticed between ‘*alam* “flag” and the entity it represents. This constituentual motivation has rendered the idiom isomorphic.

As to the examples that can be used as isomorphic or non-isomorphic, this phenomenon could be explained by the way in which (different) native speakers perceive the same idiom. In the cases discussed in (72) “tugging and pulling” and (73) “he put *sb/sth* on the shelf”, it appears that the idiomatic meanings can be perceived either as figurative abstractions that are based on entailments of the literal readings, in which case the expressions become non-isomorphic, e.g. in (72b) and (73b), or as abstractions that have semantic structures which are identical to those of the literal meanings themselves, in which case the expressions become isomorphic, e.g. in (72c) and (73c).

Partial isomorphism of the type observed in (60) “a needle in a haystack” can be attributed to the fact that the meaning “hard to locate” is an institutionalized implication of the literal reading of the expression. This meaning part is implied by what we know about the physical characteristics of both needles and haystacks. It is not part of the compositional meaning of the expression, and, therefore, does

not have a place in a structured semantic representation that is based only on this meaning.

Besides, in this example, and in others too, some implications that are communicated by the idiom, e.g. the perplexity that someone who actually looks for a needle in a haystack could experience, are not part of the literal reading of the idiom and do not have a place in a semantic representation that is based solely on it. This shows how difficult it is to paraphrase idioms literally without losing parts of their meanings. Indeed, this phenomenon reflects one of the communicative functions of metaphors in general, that is, their ability “to communicate complex configurations of information” in a compact way of expression (Gibbs 1994: 125). Finally, examples such as (60), (72), and (73) demonstrate how difficult it is in some cases to set a clear-cut distinction between metaphor and metonymy.

4 The discursive behavior of Arabic idioms

This chapter examines how Arabic idioms contribute to the content and structure of their texts. It contains two major sections. The first section looks at the discursive functions that idioms perform both in terms of what typical functions may be linked to different idioms and how contextual factors might affect this linkage in actual use. The second section concentrates on the major ways in which idioms contribute to the cohesion of their texts.

In this work, the term *text* is used to refer to a particular stretch of language, written or spoken, that is complete in its own right, e.g. a newspaper article. The term *discourse* is taken to mean a text in its situational and socio-cultural context (see Stubbs 1983 and Moon 1998).

4.1 Discursive functions of idioms

This section looks at what native speakers of Arabic do with idioms. In more technical parlance, it is concerned with the functions that these expressions have in discourse. Idioms contribute different types of content to their texts, and an awareness of the nature of these types is crucial for successful communication. Moon (1998) emphasizes that the roles that multiword units in general “have in real-time discourse are no less significant than their lexical, syntactic, and semantic characteristics. Neglecting or ignoring these roles may lead to discursal ill-formedness in encoding and to misinterpretation in decoding” (p. 219).

Nunberg *et al.* (1994) point out that studying how idioms function in discourse and what figurative patterns underlie them plays a major role in explaining their formal and semantic behavior. They also criticize some accounts of idioms and present their own approach as an advantageous alternative to these:

The problem with these accounts, on our view, is that they have tended to overgrammaticize the phenomena – to ask the syntactic or semantic apparatus of the grammar to explain regularities that are in fact the consequences of independent rhetorical and discursive functions of the expressions. By contrast, we will argue here that in order to explain the properties of these expressions fully, we have to appeal not just to the semantic properties of

idioms, but to the figurational processes that underlie them and the discursive functions that they generally serve.

(Nunberg *et al.* 1994: 494)

Indeed, examining the discursive functions of idioms may help explain different aspects of their formal behavior. For example, Strässler (1982) found that verbal idioms that express negative evaluations of their subjects tend not to occur in the first person form.

In the following analysis, I have made use of the ideas and categories discussed particularly in Moon (1998), Thompson and Hunston (2001), Fernando (1996), and Channell (2001). The labels deployed here are mainly based on Moon (1998). However, my characterization of some of them may differ to some extent from hers. This is primarily because I have attempted to incorporate insights from other relevant works as well.

The functional classes discussed below are based on analyzing how Arabic idioms contribute to the content and structure of their texts. The specific discursive function of an idiom is determined by establishing the type of contribution that it usually communicates in the contexts in which it occurs. Four basic functional categories have been identified for the idioms in the data. The following is a list of these along with some comments on what they refer to. More details on these categories and some of their examples in the data are provided in later sections:

- *Informational idioms*: these refer to entities, actions, states, or qualities.
- *Evaluative idioms*: these express the speaker's or writer's attitude towards one or more discourse referents.
- *Modalizing idioms*: these convey different meanings that are related to modality.
- *Organizational idioms*: these signal the structure of the text or indicate different types of relationship between its referents and propositions.

Some preliminary points are in order. First, the data analysis shows that idioms very often perform two or more functions simultaneously. For example, some verbal idioms that communicate actions express as well positive or negative evaluations of their subjects. Second, it has become evident from the analysis that the precise contribution of an idiom in a specific text is contingent on context. For instance, an idiom that usually conveys a negative evaluation might express a positive evaluation in some of its corpus examples, depending on how relevant contextual factors are managed. In this regard, idioms behave in a manner similar to other types of lexical items.

Third, some of the analyses of the evaluative ingredient in idioms below have strong links to the notion of semantic prosody. This refers to the phenomenon that a given word or expression with an informational or modalizing content may collocate very often with words or expressions that convey positive or negative evaluations. Consequently, this word or expression becomes associated with this

positive or negative evaluation, and this linkage can be utilized by speakers and writers to express their evaluations less openly (see Louw 2000 and Channell 2001). An example of this in English is the phrasal verb *set in*, which is often used with subjects that refer (in context) to unpleasant things, e.g. despair, infection, and rain (see e.g. Xiao and McEnery 2006).

Finally, as noted in Chapter 2, I sometimes have consulted other data sources in addition to *Arabicorpus* in order to establish the uses of the idioms that do not occur frequently in its texts. However, my searches show that some meanings of polysemous idioms, and indeed some idioms, do not occur frequently in the sources I draw on. As a consequence, consulting these sources in such cases has not provided me with much help. I have therefore based my decision regarding the relevant discursive functions only on the examples available for me to examine. This might give rise to a degree of subjectivity in some of the analyses below (see Channell 2001, who demonstrates that the evaluative content of many lexical items can only be established by examining many of their corpus examples).

Another reason for this degree of subjectivity is that in some cases, ascertaining the exact contribution of the idiom needed significant knowledge of the context. This was not always easily obtainable through the corpus texts. In this regard, one may draw attention to the fact that the newspapers included in *Arabicorpus* belong to different Arab countries, and sometimes they discuss topics that I may not be (completely) familiar with. Although I have tried to obtain the knowledge needed so as to make an informed decision, there is still a chance that in some cases relevant contextual factors were not available for me to analyze.

The results of the data analysis are reported in sections 4.1.1 to 4.1.6, and the discussion of these results is provided in section 4.1.7. Section 4.1.1 focuses on the distribution of the data over the different functional categories. Sections 4.1.2 to 4.1.5 concentrate on idioms with mainly single discursive functions. Section 4.1.6 is devoted to idioms that perform different combinations of discursive functions simultaneously. This division has led this latter section to include additional examples of some of the phenomena reported in its preceding ones. For this reason, separating the results of the data analysis from their discussion has been convenient and helpful in reducing repetition.

4.1.1 *Distribution of the data over functional categories*

Information on the distribution of the data over the functional categories is given in Table 4.1. The figures mainly reflect the distribution of the idiomatic meanings over different discursive functions (see section 3.2). In this table, idioms that perform more than one function at once have been taken into account. The information below concentrates in the main on the typical roles of idioms. Some examples of how contextual factors might affect these roles are provided in later sections.

A relevant point in this context is that an idiom might be associated with more than one typical discursive function, depending on the types of lexical item it collocates with. An example of this is:

the case in (1b) where the action which the idiom denotes and the discourse referents relevant to it are negatively evaluated. In this particular context, these are the American government and those who acted in accordance with its policies. In general, the involvement of the indefinite relative pronoun *man* “that who” which is inherently vague (see Badawi *et al.* 2004: 736) helps conceal the identities of those on whom the writer passes their negative judgment.

In (1c), the relative pronoun refers generally to types of vehicle that are similar to cars and trucks. In this case, the idiom is only informational, that is, it does not negatively or positively evaluate any referent in discourse. Considering the behavior of this idiom in the data, one may conclude that when the relative pronoun refers to a person, country, or ideology, the idiom communicates a negative evaluation in addition to its informational content. But it does not perform this evaluative function when the relative pronoun refers to an object.

4.1.2 *Informational idioms*

Informational idioms contribute to their texts by referring to entities or expressing actions, qualities, or states. The type of information conveyed can usually be linked to the syntactic class of the idiom. For instance, denoting entities is typically the function of nominal idioms, e.g.:

- (2) tuffāḥat-u ʾādam
apple-NOM Adam
Adam’s apple

Expressing states and actions is usually the function of verb-subject and verbal idioms, e.g.:

- (3) a. *shay*_i ʾakhadha baʿḍ-u-*hu*_i bi-riqāb-i baʿḍ
*sth*_i hold.on.PST some.of-NOM-POSS_i to-neck.PL-GEN some
*sth*_i (e.g. a set of ideas, texts, or problems) interconnected
- b. *huwa* majmūʿat-u nuṣūṣ-i-n yaʾkudh-u
3SG.M group.F-NOM text.PL-GEN-INDF hold.on.PRS-IND
baʿḍ-u-hā bi-riqāb-i baʿḍ (Al-Thawra, after 2000)
some.of-NOM-F.3SG.POSS to-neck.PL-GEN some
lit. it is a group of texts that hold on to each other’s necks
it is a group of texts that are connected with each other
- (4) a. *shadda* l-ḥizām-a
tighten.PST DEF-belt-ACC
he tightened his belt
- b. *daʿa* l-masʾūl-īna ʾilā *shadd-i* l-ʾaḥzimat-i
invite.PST.M.3SG DEF-official-PL.ACC to tightening-GEN DEF-belt.PL-GEN
wa-l-muḍiyy-i fī bināʾ-i l-irāq (Al-Hayat, 1997)
and-DEF-continuing-GEN in building-GEN Iraq
he invited the officials to tighten the belt and continue to build Iraq

However, idioms belonging to other syntactic classes may communicate some of these general types of meaning too. For instance, the following prepositional idiom conveys the state of being offered for discussion:

- (5) ‘ala l-ṭāwilat-i
on DEF-table-GEN
on the table

4.1.3 Evaluative idioms

The sole function of evaluative idioms is to express the speaker’s or writer’s attitude towards one or more discourse referents. Two main parameters of evaluation can be distinguished in the data: the good-bad parameter and the important-unimportant one (see Thompson and Hunston 2001). Only two idioms in the sample are purely evaluative. Both convey evaluations of a positive orientation, i.e. good or important:

- (6) a. wāsiṭat-u l-‘iqd-i
middle.part-NOM DEF-necklace-GEN
the most important or best part of something good
- b. kāna-t-i l-muḥāḍarat-u wāsiṭat-a l-‘iqd-i
be.PST-F DEF-lecture-NOM middle.part-ACC DEF-necklace-GEN
fī barnāmaj-i-n ḥāfil-i-n (Al-Hayat, 1996)
in program-GEN-INDF varied-GEN-INDF
lit. the lecture was the middle jewel of the necklace in a varied program
the lecture was the best part in a varied program
- (7) a. marbaṭ-u l-faras-i
place.where.sth.is.tied.up-NOM DEF-mare-GEN
the most important point (e.g. in a problem or an argument)
- b. fa-l-muwāṣafāt-u ’idhan wa-l-iltizām-u bi-hā
and.so-DEF-standard.PL.F-NOM then and-DEF-adhering-NOM to-F.3PL
hiya marbaṭ-u l-faras-i fī ’amaliyyāt-i
3PL.F place.where.sth.is.tied.up-NOM DEF-mare-GEN in activity.PL-GEN
l-taṣḍīr-i (Al-Ahram, 1999)
DEF-exporting-GEN
lit. standards and adhering to them then are the place where the mare is tied up in exporting activities
lit. standards and adhering to them then are the most important factor in exporting activities

In (6b), the positive evaluation relates to both the lecture and the varied program. In (7b), the idiom positively evaluates adhering to the standards in exporting activities.

Relevant changes in the context may affect the evaluative content of the idiom. As Fernando (1996) notes, “[a]typical appraisals arise from appropriate management of the co-text accompanying the idiom” (p. 146). Moon (1998: 254) also points out that:

[t]he canonical evaluative orientation of FEIs [i.e. fixed expressions and idioms] cannot always be taken for granted, and corpus evidence shows that in context evaluations are sometimes modified in the co-text or framing narrative, perhaps through the addition of a negative or modal. In particular, FEIs denoting actions or situations that are conventionally and culturally marked as good or bad may be used in contexts where their polarity is reversed and the evaluation renegotiated instantially.

In the following corpus example, the idiom in (6a) does not express the same degree of positive evaluation it typically does:

- (8) taḥawwala-t-i l-madīnat-u llatī kāna-t bi-lā qalb-i-n
 transform.PST-F DEF-city.F-NOM REL be.PST-F with-NEG heart-GEN-INDF
 fī shi‘r-i ḥijāzī ‘ilā wāsiṭat-i ‘iqd-i-n
 in poetry-GEN Hijazi to middle.part-GEN necklace.M-GEN-INDF
 mutamāthil-i l-ḥabbāt-i yaḍumm-u kull-a l-mudun-i
 identical-GEN DEF-pearl.PL-GEN include-IND all-ACC DEF-city.PL-GEN
 l-‘arabiyyat-i (Al-Hayat, 1996)
 DEF-Arab-GEN
 lit. the city, that was heartless in Hijazi’s poetry, transformed to the
 middle jewel of a necklace of identical pearls that includes all the Arab
 cities

In this context, the idiom does not portray the entity that it evaluates, here the city, as the best, but only as one of equally good or important cities. In other words, although the idiom still positively evaluates the city, the degree of this evaluation has been reduced by manipulating the original metaphorical image.

4.1.4 *Organizational idioms*

Organizational idioms in the data signal the temporal, spatial, or logical relations between the propositions or discourse referents (see Moon 1998: 234 and Fernando 1996). Examples of these include:

- (9) a. ‘alā marmā ḥajar-i-n
 at distance.that.sth.is.thrown stone-GEN-INDF
 of proximity in space or time
 b. yaqa‘-u l-mathaf-u l-’ūlimbiyy-u l-shaḥīr-u
 be.located-IND DEF-museum-NOM DEF-Olympian-NOM DEF-famous-NOM
 ‘alā marmā ḥajar-i-n min-a
 at distance.that.sth.is.thrown stone-GEN-INDF from
 l-funduq-i (Al-Hayat, 1997)
 DEF-hotel-GEN
 the famous Olympian Museum is located within a stone’s throw from the
 hotel

- (10) a. fī 'a'qāb-i shakhṣ/shay'
 in heel.PL-GEN sb/sth
 close behind *sb*, or immediately after *sb/sth*
- b. 'ashāra 'ila 'anna mīlūševič tarāja'a fī 'uktūbar fī 'a'qāb-i
 refer.PST to COMP Milošević retreat.PST in October in heel.PL-GEN
 l-tahdīd-i bi-shann-i ḥarb-i-n
 DEF-threat-GEN with-launching-GEN war-GEN-INDF
 jawwiyyat-i-n (Al-Ahram, 1999)
 airy-GEN-INDF
 lit. he pointed out that Milošević retreated in October in the heels of the
 threat to launch an airy war
 he pointed out that Milošević retreated in October after the threat to
 launch an air war
- (11) a. fī ḍaw'-i shay'
 in light-GEN sth
 in the light of *sth*
- b. fī ḍaw'-i mā sabaqa yanbaghī li-kull-i dawlat-i-n
 in light-GEN REL precede.PST should for-all-GEN country.F-GEN-INDF
 'alā ḥīdat-i-n 'an takhtār-a niẓām-a 'as'ār-i
 on solitude-GEN-INDF COMP choose-SBJV regime-ACC rate.PL-GEN
 l-ṣarf-i lladhī yakūn-u 'akthar-a
 DEF-money.changing-GEN REL be.PRS-IND more-ACC
 mulā'amat-a-n li-'awḍā'-i-ha
 suitability-ACC-INDF for-situation.PL-GEN-F.3SG.POSS
 l-iqtisādiyyat-i wa-l-māliyyat-i (Al-Hayat, 1997)
 DEF-economic-GEN and-DEF-financial-GEN
 in the light of what preceded, every country should choose by itself the
 exchange rate regime that is more suitable for its economic and financial
 situation

4.1.5 Modalizing idioms

Modalizing idioms express the degree of the reality of the proposition or phenomenon as perceived by the speaker or writer (see Trask 1993). Some meanings that belong to modality are notions of probability, improbability, certainty, doubt, ignorance, expressions of vagueness, retreating from definite assertion, indications of the extent or degree of the phenomenon, and indications of the range of the reference by pointing to the degree of inclusiveness (see Palmer 1990 and Moon 1998).

The following modalizing idioms express the degree of inclusiveness:

- (12) a. *shakhṣ/shay'*_i bi-rummat-i-*hi*_i
*sb/sth*_i with-worn.out.piece.of.rope-GEN-M.3SG.POSS_i
 all *sb/sth*_i
- b. qad tu'arriḍ-u l-niẓām-a l-'idāriyy-a
 may expose-IND DEF-system.M-ACC DEF-administrative.M-ACC

4.1.6.1 Non-evaluative multifunctional idioms

These include only informational-modalizing idioms which both contribute to the informational content of their texts and incorporate a modality ingredient, e.g.:

- (14) a. qarṣat-u ʾudhun-i-n
 pinch-NOM ear-GEN-INDF
 a light non-physical punishment that serves as a warning of a possible cruel one
- b. ʾinna khasārat-a l-mustaqillat-i bi-l-ʾulūm-i
 indeed defeat.F-ACC almustaqillat.F-GEN in-DEF-science.PL-GEN
 l-ijtimāʿiyyat-i qad takūn-u ʾibārat-a-n ʾan
 DEF-social-GEN may be.PRS.F-IND equivalent-ACC-INDF to
 qarṣat-i ʾudhun-i-n li-qiyādiyy-ī-hā (Al-Watan, 2002)
 pinch-GEN ear-GEN-INDF for-leader-PL.GEN-3SG.F
 lit. indeed, the defeat of the Al-Mustaqillah [list] in [the student union elections in the faculty of] Social Sciences may be a pinch of an ear for its leaders
 indeed, the defeat of the Al-Mustaqillah list in the student union elections in the faculty of Social Sciences may be a warning for its leaders
- (15) a. taḥta l-ʾaḍwāʾ-i
 under DEF-light.PL-GEN
 in the spotlight
- b. ʾinna wafd-a l-dawlat-i l-ʾibriyyat-i
 indeed delegation.M-ACC DEF-state-GEN DEF-Hebraic-GEN
 kāna taḥta l-ʾaḍwāʾ-i wa-markaz-a
 be.PST.M.3SG under DEF-light.PL-GEN and-center-ACC
 jadwal-i l-ʾaʿmāl-i (Al-Hayat, 1996)
 schedule-GEN DEF-work.PL-GEN
 lit. indeed, the delegation of the Hebraic state was under the lights and the center of the schedule of works
 indeed, the Hebraic state's delegation was in the spotlight and the center of the agenda
- (16) a. lā yuṣaddiq-u ʾayn-ay-*hi*
 NEG believe.M.3SG-IND eye-DU.ACC-M.3SG.POSS
 he does not believe his eyes
- b. lam nuṣaddiq ʾaʿyun-a-nā ʾindamā
 NEG.PST believe.1PL.JUSS eye.PL-ACC-1PL.POSS when
 dakhala l-sayyid-u l-raʾīs-u
 enter.M.3SG.PST DEF-mister-NOM DEF-president-NOM
 wa-l-sayyidat-u ʾaqīlat-u-hu li-tanāwul-i
 and-DEF-lady-NOM wife-NOM-M.3SG.POSS for-having-GEN

l-'iftār-i ma'a-nā (Al-Thawra, after 2000)
 DEF-breakfast-GEN with-1PL
 we did not believe our eyes when Mr. President and his wife came in to
 have breakfast with us

The idiom in (14b) both refers to the phenomenon, i.e. the punishment, and indicates its extent by situating it on the harsh-soft scale, here it is only a light punishment. The idiom in (15b) refers to the state of receiving public attention and indicates the high intensity of this attention at the same time. Finally, in (16b), the idiom both refers to the state of being surprised and indicates the great extent of this feeling.

4.1.6.2 *Evaluative multifunctional idioms*

A key point that features in the following analysis of the evaluative ingredient in idioms is the element or elements to which the evaluation relates, that is, the action, quality, situation, and/or entity on which the speaker or writer passes their judgment.

The three subtypes belonging to evaluative multifunctional idioms comprise 68 cases. Nineteen of these are typically associated with positive evaluations, and 42 with negative ones. In other cases, the orientation of the evaluation can only be decided on in context. Finally, some idioms convey both positive and negative evaluations that relate to different discourse referents.

4.1.6.2.1 INFORMATIONAL-EVALUATIVE IDIOMS

These idioms denote e.g. actions, situations, or qualities, and, at the same time, express the speaker's or writer's attitude. Some examples of this type in the data are provided below and followed by some analytical comments. Since the informational content of each of these examples is stated explicitly in the English translation provided for it, the comments mainly concentrate on the orientation of the evaluation and the element(s) to which it relates in context:

- (17) a. ḥibr-u-n 'alā waraq-i-n
 ink-NOM-INDF on paper-GEN-INDF
 inactive (e.g. of decisions or laws)
- b. lākin wa-lil'asaf baqiya l-qarār-u
 but and-unfortunately remain.PST DEF-decision-NOM
 ḥibr-a-n 'alā waraq-i-n (Al-Thawra, after 2000)
 ink-ACC-INDF on paper-GEN-INDF
 lit. but, unfortunately, the decision remained ink on paper
 but, unfortunately, the decision remained inactive
- (18) a. ištāda fi l-mā'-i l-'akir-i
 fish.PST in DEF-water-GEN DEF-cloudy-GEN
 he fished in troubled waters

- b. huwa yushīr-u 'ilā naw'īyyat-i-n min-a l-nās-i
M.3SG refer.PRS-IND to type.F-GEN-INDF of DEF-people-GEN
'tāda-t 'an taṣṭād-a fi l-mā'-i l-'akir-i
use.PST-F COMP fish.PRS-SBJV in DEF-water-GEN DEF-cloudy-GEN
wa-'tāda-t 'an tastaḥīd-a min maṣā'ib-i
and-use.PST-F COMP benefit.PRS-SBJV from disaster.PL-GEN
ghayr-i-hā (Al-Hayat, 1997)
other.than-GEN-F.3SG
he refers to a type of people that used to fish in troubled waters and gain advantages from other people's disasters
- (19) a. yanfukh-u fī qirbat-i-n maqtū'at-i-n
blow.PRS-IND into bagpipe-GEN-INDF cut.PPTCP-GEN-INDF
he is trying in vain to convince other people to improve a situation
- b. sāmiḥ-ū-nā 'idhā 'aṣrarnā 'alā muwāṣalat-i
forgive.IMP-2PL-1PL COND insist.PST.1PL on continuing-GEN
l-nafkh-i fi l-qirbat-i l-maqtū'at-i
DEF-blowing-GEN into DEF-bagpipe-GEN DEF-cut.PPTCP-GEN
wa-waqaf-nā fī makān-i-nā nuḥāwil-u radd-a
and-stand.PST-1PL in position-GEN-1PL.POSS try.PRS.1PL-IND regaining-ACC
l-i'tibār-i wa-raḥ-a l-mahānat-i 'an
DEF-respect-GEN and-removing-ACC DEF-humiliation-GEN away.from
lughat-i-na l-waṭaniyyat-i (Al-Ahram, 1999)
language-GEN-1PL.POSS DEF-national-GEN
lit. forgive us if we insist on continuing blowing into the defective bagpipe and if we stand in our position trying to regain respect for and remove the humiliation away from our national language
- (20) a. nazīf-u l-yad-i
clean-NOM DEF-hand-GEN
acting ethically
- b. wa-'anta 'ayyuha l-qā'id-u l-nazīf-u l-yad-i
and-M.2SG VOC DEF-leader-NOM DEF-clean-NOM DEF-hand-GEN
l-ṭāhir-u l-nafs-i inṭaliq bi-nā 'ilā
DEF-pure-NOM DEF-soul-GEN go.quickly.IMP.2SG with-1PL to
'āfāq-i mustaqbal-i-n mushriq-i-n (Al-Ahram, 1999)
horizon.PL-GEN future-GEN-INDF bright-GEN-INDF
lit. and you, O, the clean-handed, pure-soul leader, lead us to the horizons of a bright future
- (21) a. ghuṣn-u l-zaytūn-i
branch-NOM DEF-olive-GEN
olive branch
- b. yuqaddim-u ghuṣn-a l-zaytūn-i li-hā'ulā'i l-ladhīna
present.PRS.M.3SG-IND branch-ACC DEF-olive-GEN to-DEM.PL REL
khāṣam-u l-silm-a (Al-Ahram, 1999)
set.oneself.against-3PL DEF-peace-ACC
lit. he presents the olive branch to those who set themselves against peace

In (17b), the idiom disapproves of inactivity. But, in some examples, the contexts may also suggest a negative evaluation of the entity that is deemed responsible for not bringing e.g. the decision, law, or agreement into force. The idiom in (18b) disapproves of both gaining advantages from other people's problems and the type of people who used to do so. The idiom in (19b) negatively evaluates the situation that the subject wants to make better. Sometimes, it also suggests disapproval of those who the subject is trying to convince to take the actions needed to improve the situation. In this particular corpus example, the context shows that the negatively evaluated entity is those who support the use of English as the language of instruction in the faculties of science in Egyptian universities. It may be added too that the subject of this idiom is often regarded positively and considered, as some native Arabic speakers point out, as a "reformer".

The idioms in (20a) and (21a) convey positive evaluations. In (20b), the idiom positively evaluates acting according to ethics and the leader doing this. However, at least in some cases, the mere reference to and praise of this quality may induce a negative evaluation of those who do not show it. In (21b), the positive evaluation mainly relates to the notion of peace. However, in many cases, when this idiom is used, any discourse referent that accepts or offers peace is positively evaluated, and any discourse referent that rejects it is negatively evaluated.

Some other examples of informational-evaluative idioms are given below without analytical comments. Typically, the idioms in (22) to (28) convey negative evaluations, whereas the idioms in (29) to (31) express positive evaluations:

- (22) 'adāra l-khadd-a l-'ākhar-a
turn.PST DEF-cheek-ACC DEF-other-ACC
he turned the other cheek
- (23) bala'a l-ṭu'm-a
swallow.PST DEF-bait-ACC
he swallowed the bait
- (24) baqarat-u-n ḥalūb-u-n
cow-NOM-INDF milch-NOM-INDF
milch cow
- (25) dam-u-n 'azraq-u
blood-NOM-INDF blue-NOM
blue blood
- (26) 'ustūwānat-u-n mashrūkh-at-u-n
record.F-NOM-INDF crack.PPTCP-F-NOM-INDF
repetitious boring talk about a particular topic
- (27) bayna nār-ayni
between fire-DU.GEN
between the devil and the deep blue sea
- (28) min warā'-i zahr-i shakhṣ
from behind-GEN back-GEN sb
behind sb's back

- (29) ʾakmala nişf-a dīn-i-hi
 perfect.M.3SG.PST half-ACC religion-GEN-M.3SG.POSS
 he got married
- (30) ɗaraba ʿuşfūr-ayni bi-ḥajar-i-n
 hit.PST sparrow-DU.ACC.INDF with-stone-GEN-INDF
 he killed two birds with one stone
- (31) wāsiʿ-*u* l-ʾufuq-i
 broad-NOM DEF-horizon-GEN
 broad-minded

Context plays an important part in establishing the elements to which the evaluative content relates. For example, the idiom in (32a) is often used to refer to non-human entities that are used as ready sources of easy profit. The negative evaluation typically relates to whoever takes advantage of the entity that the idiom refers to in context, but not to the entity itself, as in (32b). However, in other contexts, e.g. in (32c), when this idiom is used to refer to a person, it could suggest some negative connotations of him or her, since it would depict a human being in terms of an animal. Therefore, the speaker or writer should be cautious if they do not want this to happen, as (32c) shows:

- (32) a. baqarat-u-n ḥalūb-u-n
 cow-NOM-INDF milch-NOM-INDF
 milch cow
- b. wa-bi-sabab-i baʿd-i l-ḥukkām-i lladhīna jaʿal-ū
 and-by-reason-GEN some-GEN DEF-ruler.PL-GEN REL.PL make.PST-3PL
 min-a l-istiqlāl-i baqarat-a-n ḥalūb-a-n
 out.of DEF-independence-GEN cow-ACC-INDF milch-ACC-INDF
 (Al-Ahram 1999)
 lit. and because of some rulers who made a milch cow out of independence
 and because of some rulers who used independence as a milch cow
- c. wa-ʾaşbaḥa baʿd-u ʾaşḥāb-i l-munshaʾāt-i
 and-become.PST some-NOM owner.PL-GEN DEF-establishment.PL-GEN
 maʿa l-iḥtirām-i la-hum baqarat-a-n ḥalūb-a-n
 with DEF-respect-GEN for-M.3PL cow-ACC-INDF milch-ACC-INDF
 (Al-Thawra, after 2000)
 and some establishment owners have become, with all due respect, a milch cow

In its original text, the Arabic idiom in (32c) is typed within two forward slashes with no apparent case endings. This might be explained by one or both of the following reasons. First, the slashes might be employed here to signal the use of non-literal language. Second, this expression is sometimes used in spoken Arabic. Therefore, using it between forward slashes without case endings could be taken as an indication that the expression is somewhat

colloquial, at least from the viewpoint of the writer. However, it should be noted that neither of these seems to be a standard use of this punctuation mark in Arabic today.

Another interesting example of the interaction between the idiomatic meaning and the contexts in which the idiom is used is (33a). Although this idiom refers to nobility, in the data it is often used derogatorily in contexts that disapprove of the group of people that is considered of a higher rank or class, as in (33b):

- (33) a. dam-u-n ’azraq-u
 blood-NOM-INDF blue-NOM
 blue blood
- b. al-fard-u l-muṭlaq-u yubarrir-u
 DEF-one-NOM DEF-absolute-NOM justify.PRS-IND
 sulṭān-a-hu bi-hādhīhi l-’asāṭīr-i llatī
 authority-ACC-M.3SG.POSS with-DEM DEF-myth.F.PL-GEN REL
 yarwī-ha l-balāṭ-u ‘an dam-i-hi
 recount.PRS.IND-F.3PL DEF-court-NOM about blood-GEN-M.3SG.POSS
 l-nabīl-i l-’azraq-i (Al-Hayat, 1997)
 DEF-noble-GEN DEF-blue-GEN
 the absolute ruler justifies his authority with those myths that the court
 recounts about his noble blue blood

At least in some of its examples, this idiom also suggests disapproval of the notion and practice of discrimination on the basis of race, color, etc. in general.

The orientation of the evaluative content of some idioms can only be decided on in context, e.g.:

- (34) a. rafa‘a rāyat-a shay’
 raise.PST flag-ACC sth
 he supported sth
- b. tarfā‘-u rāyat-a l-waḥdat-i
 raise.PRS.F.3SG-IND flag-ACC DEF-unity-GEN
 l-‘arabīyyat-i (Al-Hayat, 1996)
 DEF-Arab-GEN
 lit. it raises the flag of the Arab unity
 it supports the Arab unity
- c. rafa‘-ū rāyat-a l-shayṭān-i (Al-Tajdid, 2002)
 raise.PST-3PL flag-ACC DEF-devil-GEN
 lit. they raised the flag of the devil

In (34b), the context shows that the idiom conveys a positive evaluation of the action and the subject, here the MBC TV station. In (34c), the context shows that the action and the subject, i.e. the imperial powers, are negatively evaluated.

Finally, some idioms convey both positive and negative evaluations of different referents in their contexts. For instance, the idiom in (35a) typically, but not always, occurs in contexts where the subject of the verb is negatively evaluated and the thing prevented is positively evaluated, as in (35b):

- (35) a. waqafa fī ṭarīq-i *shakhṣ/shay'*
stand.PST in way-GEN *sb/sth*
he stood in the way of *sb/sth*
- b. yaqif-a l-jāmid-ūna wa-l-muḥāfiẓ-ūna
stand.PRS-SBJV DEF-rigid-PL.NOM and-DEF-conservative-PL.NOM
fī ṭarīq-i tayyār-i l-'iṣlāḥ-āt-i
in way-GEN current-GEN DEF-reform-F.PL-GEN
l-ijtimā'īyy-at-i (Al-Ahram, 1999)
DEF-social-F-GEN
lit. the rigid and the conservative stand in the way of the current of social reforms

4.1.6.2.2 EVALUATIVE-MODALIZING IDIOMS

These idioms convey an evaluation of a phenomenon, situation, or entity and also communicate the speaker's or writer's estimation of its degree of reality, e.g.:

- (36) a. qimmat-u jabal-i l-jalīd-i
top-NOM mountain-GEN DEF-ice-GEN
the tip of the iceberg
- b. al-ḥālāt-u llatī tamma l-kashf-u 'an-hā
DEF-case.F.PL-NOM REL be.done.PST DEF-disclosure-NOM of-F.3PL
hatta l'āna laysa-t siwā qimmat-i jabal-i
until now NEG-F but top-GEN mountain-GEN
l-thalj-i (Al-Ahram, 1999)
DEF-ice-GEN
lit. the cases that have been disclosed until now are not but the summit of the ice mountain
the cases that have been disclosed until now are not but the tip of the iceberg

Here, the cases that have been disclosed are assessed as only a small part of a much bigger problem (see Moon 1998: 240, for a relevant analysis of the English idiom *the tip of the iceberg*).

4.1.6.2.3 INFORMATIONAL-EVALUATIVE-MODALIZING IDIOMS

Idioms belonging to this category contribute informationally to their texts and also include both evaluative and modalizing components. Some examples of these

are given below. In (39b), *that and which* is used in the literal rendition of the corpus example only as a convenient translation of the idiom:

- (37) a. 'arghā wa-'azbada
foam.PST and-froth.PST
he fumed with rage
- b. waqafa-t ma'ī 'umm-ī wa-'ikhwat-ī
stand.PST-F with.1SG mother-1SG.POSS and-brother.PL-1SG.POSS
wa-'ikhwat-u-hu lākinna-hu 'arghā
and-brother.PL-NOM-M.3SG.POSS but-M.3SG foam.PST.M.3SG
wa-'azbada wa-qāṭa'a 1-jamī'a
and-froth. PST.M.3SG and-break.with.PST DEF-everyone-ACC
wa-'ajbara-nī 'alā muqāṭa'at-i-him (Al-Ahram, 1999)
and-force.PST.M.3SG-1SG on breaking.with-GEN-3PL
lit. my mother stood by me and so did my brothers and his brothers, but
he foamed and frothed and broke with everyone and forced me to break
with them
my mother stood by me and so did my brothers and his brothers, but
he fumed with rage and broke with everyone and forced me to break
with them
- (38) a. dhāqa marārat-a *shay'*
taste.PST bitterness-ACC *sth*
he experienced the suffering of *sth*
- b. dhāqa marārat-a 1-hazīmat-i (Al-Ahram, 1999)
taste.PST bitterness-ACC DEF-defeat-GEN
lit. he tasted the bitterness of defeat
- (39) a. allutayyā wa-llatī
REL.F.SG.DIMIN and-REL.F.SG
much ado
- b. fa-stahwā-na 1-kitāb-u wa-badhal-nā ghāyat-a
and.so-attract-1PL DEF-book-NOM and-make.PST-1PL utmost-ACC
1-majhūd-i li-l-ḥuṣul-i 'alayhi wa-ba'da
DEF-effort-GEN for-DEF-obtaining-GEN on.3SG.M and-after
llutayyā wa-llatī kamā yuqāl-u wašala 'ilayna
REL.F.SG.DIMIN and-REL.F.SG as say.PASS-IND arrive.PST to.1PL
1-kitāb-u (Al-Hayat, 1996)
DEF-book-NOM
lit. and so the book attracted us, and we made the utmost effort to obtain
it, and after that and which, as it is said, the book arrived to us
so the book attracted our interest and we made the utmost effort to obtain
it, and, after much ado, as they say, it arrived
- (40) a. nār-u-n 'alā 'alam-i-n
fire-NOM-INDF on mountain-GEN-INDF
very well-known entity

actions, events, situations, people and things, attributes, and evaluations (Fernando 1996). In her characterization of this category, Fernando (1996) points out that the Hallidial term *ideational* is used:

to designate the macro-function of language realized through the clause and concerned with articulating the speaker's or writer's experience of the world: participants, actions and processes, the attributes of the participants and the circumstances associated with actions and processes [. . .]. Typically, ideational idioms are realized by units smaller than the clause [. . .]. These units function as parts of clauses (pp. 97–98).

Therefore, it may be said that, once more, regarding idioms as only fully non-compositional multiword units that occur within the clause may provide an explanation for the frequency of these two functions in the data.

In addition, the high frequency of the informational function in particular can be attributed simply to the fact that giving information is an essential function of language (Moon 1998).

A substantial number of the idioms in the data convey evaluations. Similar observations have been made in previous work. For instance, Nunberg *et al.* (1994: 493) note that:

[i]dioms are typically used to imply a certain evaluation or affective stance toward the things they denote. A language doesn't ordinarily use idioms to describe situations that are regarded neutrally – buying tickets, reading a book – though of course one could imagine a community in which such activities were sufficiently charged with social meaning to be worthy of idiomatic reference.

Carter (2004: 132) also points out that:

idioms are not simply neutral alternatives to less semantically opaque expressions. There is a difference between “I smell a rat” and “I am suspicious”, or “She's on cloud nine” and “She's extremely happy”, or “The garden's a real mish-mash of different herbs” or “The garden's a real mixture of different herbs.” In all cases the idiomatic expression is used evaluatively and represents a more intense version of the literal statement.

The evaluative ingredient in idioms may “arise from their informational content, the imagery conveying this content, habitual collocates, and habitual contexts of use over time and space” (Fernando 1996: 146). Examining the first two sources on the basis of synchronic data can help establish why many idioms have evaluative functions. On the other hand, examining the habitual collocates of an idiom using only synchronic data may show what evaluative content the idiom has. But it may not necessarily explain how the idiom has acquired this content in the first place, since it is possible that any patterns in the idiom's collocability are in fact a

result of the evaluative content it acquired at an earlier stage of its life rather than the trigger of this content. Indeed, looking at the usual collocates of idioms has often helped me establish their evaluative content. Finally, investigating the habitual contexts of use over time and space would require a diachronic approach, which is not taken in this study. Therefore, any explanations suggested below that rely (partly) on this diachronic dimension are only given as hypotheses that need to be tested in future work.

The informational content of the expressions may account for the presence of the evaluative ingredient in idioms such as (18a) “he fished in the cloudy water”, (20a) “clean-handed”, (21a) “olive branch”, (31) “broad-horizoned”, and (39a) where two relative pronouns are coordinated to denote “much ado”.

Cultural norms may provide explanations for the particular orientation of the evaluation in such cases. For instance, the idiom in (18a) “he fished in the cloudy water” negatively evaluates the action it denotes and the agent. The motivation for this idiom consists in the perception that a fish would bite more readily when the water is stained, since, under these circumstances, it would be able to see the bait but not the fisherman, as one of my informants has pointed out. This motivation does not seem to be clear to many native speakers, however. Therefore, it may not explain the negative evaluative content. In fact, this motivation could be a source of a positive evaluation, since it implies skill in choosing the appropriate circumstances in which one can easily get what he or she wants. Rather, the negative orientation of the evaluative content would be better accounted for by the fact that people normally disapprove of attempts to gain advantages from difficult situations, in particular, when some people try to take advantage of others’ problems.

Cultural values may be the only source of explanation for the evaluative orientation in idioms that are not based on metaphors and/or metonymies. For example, the positive evaluative content in (21a) “olive branch” and (29) “he perfected his religion” can be attributed to the fact that the speech community highly regards peace, religion, and marriage.

Relying on the figuration underlying the idiom in (33a) “blue blood” in its non-Arabic etymological context does not appear to help in explaining its negative evaluation. In fact, it is possible that, in its original context, the expression had a positive evaluative content. Also, the typically neutral use of the English idiom *blue blood* suggests that the negative evaluation in the Arabic expression is not due to the influence of English. Thus, this negative evaluation could be ascribed to the historical developments of the contexts in which this idiom is used. That is, the idiom might have developed its negatively evaluative content because of the socio-cultural changes that have led many people to take a dim view of different sorts of discrimination on the basis of e.g. color, race, or class. Additionally, the effects of imperialism might have played a part here too.

However, a possible role of the original context cannot be completely ruled out. That is, in this context, the families to whom the idiom originally referred claimed that they were pure and had not been contaminated by people from other different races, including the Moors. Since the Moors are Muslims who are of mixed Berber and Arab descent, considering them as a source of contamination could provide a

plausible explanation for why this idiom is often used disapprovingly in Arabic. Given this analysis, it might be said that the etymological context of the idiom is not, after all, entirely irrelevant to the Arabic context. However, determining to what extent such historical circumstances actually explain the contemporary evaluative orientation would require adopting a diachronic approach to the investigation.

The property of *figurativity* that features in many Arabic idioms may provide a credible explanation for the high frequency of the evaluative function. The strong link between figurative language and evaluation has been reported in previous studies. For example, Moon (1998) found that the vast majority of the examples in her data that show “any metaphorical or simile content have some evaluative function” (p. 225). Also, Krzeszowski (1990) points out that figurative language is more likely to express evaluations than non-figurative language.

As instances of figurative language, many idioms evoke mental scenes that may not be attainable through literal language. The situations or actions that are expressed in the literal readings of idioms may be experienced as good or bad and they may be culturally regarded approvingly or disapprovingly (Moon 1998). In cognitive-linguistic terms, “[t]he literal scene works as a rich resource for weak implicatures because it encodes directly accessible basic conceptual and pre-conceptual experiences that are rich in both [*sic*] conceptual, sensory and emotional content” (Langlotz 2006: 140). These received cultural norms and implicatures may constitute the basis for the evaluative content in many idioms.

The imagery may account, fully or partially, for the presence of the evaluative content in idioms such as (6a) “the middle jewel of the necklace”, (22) “he turned the other cheek”, (27) “between two fires”, (28) “from behind *sb*’s back”, (30) “he hit two sparrows with one stone”, (36a) “the top of the ice mountain”, (37a) “he foamed and frothed”, and (38a) “he tasted the bitterness of *sth*”.

Examining the characteristics of the figuration underlying evaluative idioms, perhaps including how the literal scene is interpreted within its etymological socio-cultural context, often clarifies why they have particular orientations. As Moon (1998) points out, “[m]etaphors frequently hold the keys to the orientation of the expression, [. . .] real-world knowledge about the metaphorical image and its connotations colours and clarifies evaluation” (p. 248).

For example, in (6a) “the middle jewel of the necklace”, the positive evaluation draws on the knowledge that the middle piece of jewelry in a necklace is usually the biggest and most valuable part of it. However, a famous classical book, titled *Al-‘iqd-u l-farīd* “The unique necklace”, named its central chapter *wāsiyat-u l-‘iqd* “the middle jewel of the necklace”. This could be the source of the idiom. However, since many native speakers today may not know about this classical work, they may just motivate the idiom and understand its evaluative content only relying on interpreting its literal meaning.

In (36a) “the top of the ice mountain”, the fact that icebergs cause huge problems for ships may underlie the typically negative evaluation expressed in this idiom. In (28) “from behind *sb*’s back”, the negative evaluation may be derived from the fact that the literal scene is often associated with contexts of e.g. betrayal or cowardice, which are socially unacceptable.

The orientation of the evaluative content in idioms can be linked to the processes of *highlighting* and *hiding* that take place during the process of metaphorization. The former refers to mapping selected features from the source domain of the metaphor onto the target one. The latter refers to the suppression of some other features in the source domain (Kövecses 2002). At least in principle, some of the metaphorical images on which idioms are based can be sources of negative or positive evaluations. The specific orientation of the evaluative content depends on what aspects of the literal meaning are or are not transferred to the idiomatic one. An example of this is the idiom in (30) “he hit two sparrows with a stone”. It often conveys a positive evaluation because the element highlighted in the literal meaning is basically the skill involved in accurately targeting two aims at the same time. This meaning, however, could have been a source of a negative evaluation if cruelty to animals, which is disapproved of by many people, was highlighted.

The way in which the literal scene is interpreted within its etymological socio-cultural context could provide an explanation for the specific orientation of the evaluation in some idioms. For example, the idiom in (40a) “a fire on a mountain” often evaluates positively the well-known entity. However, the literal scene does not explain why it is used in this way. A possible explanation for this orientation is the context within which the literal scene used to occur (and maybe still occurs in certain environments). That is, setting fires on mountains at night in the desert played an important role in saving the lives of many people who were lost and/or in need of food and water. The fire guided them to a place where they were able to find what they desperately needed.

The positive way in which this scene was regarded can be noticed in classical literature. Namely, in one of her famous poems, Al-Khansā’, a celebrated classical poet, made use of a very similar image in the context of praising her brother. In this poem, she says that her brother is a shining example that people follow and uses the simile:

ka-’anna-hu ‘alam-u-n fī ra’s-i-hi nār-u-n
as.if-that-M.3SG mountain.M-NOM-INDF in head-GEN-M.3SG.POSS fire-NOM-INDF
lit. as if he is a mountain with a fire on its summit

Indeed, this literary use itself could be the source of the idiom and its positive evaluation. The fact that this expression is partly not motivated to many native speakers due to the presence of *‘alam* (see section 3.2) could support the attribution of the evaluative content solely to its historical development. However, fame in itself is highly regarded by many people today.

Also, the etymological development of the idiom in (13a) “on their father’s camel”, i.e. all together, suggests itself as an explanation for why in many cases it has a negative association. That is, the relatively frequent association of this idiom with the strong disapproval of the actions of destruction, killing, burning, and the like could be ascribed to its contexts of use that might have been affected by its (assumed) etymology, i.e. the story of the three sons who were killed and sent back carried on their father’s camel (see section 3.2.6). The fact that the motivation of

the idiom is not transparent, at least to many native speakers, could back up this as an account for the phenomenon. However, again, diachronic examinations are needed to ascertain this or not.

An important point that needs to be accounted for is that negatively evaluative idioms outnumber positively evaluative ones. Carter (2004) draws attention to the fact that “idiom uses do not simply describe but comment in positive and (more usually) negative ways on events, processes and persons” (p.132).

This phenomenon may be explained by the property of *indirectness* in idioms. As semantically non-compositional expressions, idioms are evidently indirect in communicating their meanings. This often makes them an appropriate choice when indirectness is needed in conveying certain information and/or evaluations. For instance, speakers and writers may resort to indirect language when direct messages are considered to be face-threatening or rude (Moon 1998). This also may account for the presence of many idioms that express meanings related to topics that are regarded as taboo, such as death and sex, and explain why these are frequently preferred to their literal counterparts in certain contexts (Langlotz 2006).

In general, language users need tactful ways to express negative evaluations more than they do to express positive ones. This could provide a motivation for formularizing more indirect negatively evaluating expressions than indirect positively evaluating ones. Since the use of formulaic language is characterized by the need for less cognitive processing effort (Wray 2002), formularizing many negatively evaluative non-literal expressions could then be regarded as an attempt to maximize the advantages that these expressions equip language users with.

Two advantages may be noted here. First, many idioms allude to values that are shared in the speech community. Appealing to shared values provides a shield behind which the speaker or writer can hide from critique or pre-empt disagreement when they want to communicate certain evaluations, particularly negative ones. It does this by making use of the more or less consensus about how the literal scene of the idiom is generally evaluated in the speech community. Appealing to shared values both substantiates the message conveyed and reinforces the widely accepted attitude (Moon 1998).

The other advantage of (evaluative) non-literal expressions is their ability to create and support rapport between speakers/writers and hearers/readers. This property is present in the use of indirect language in general. The skill to accurately encode and decode indirect language is often a strong indication of belonging to the same group, whereas failing to do so frequently signals outsiders. As Wray (2008) notes, “implicit encoding simultaneously supports internal social cohesion and creates a barrier to communication with outsiders who do not share the group’s knowledge” (p. 54). Gibbs (1989: 250) even raises the possibility that this could be the principal social function of metaphors. In his words:

metaphorical talk often presupposes and reinforces an intimacy between speaker and listener, the cultivation of which is, perhaps, the primary

function of such language [. . .]. Intimacy can be enjoyed by all those who are confident that what they say will be understood. It is the bond among those who share not only a basic linguistic competence, but a common stock of experiences, interests, sensibilities, and the ability to call upon that information when interpreting language.

As examples of indirect language, idioms may then help the speaker or writer with creating and/or reinforcing solidarity with their hearers or readers. In the case of some classical (opaque) Arabic idioms, one may suggest a different, though related, motivation for their selection in discourse. That is, by using them, some writers might want to show their mastery of classical Arabic (phraseology), a quality that is appreciated within some literary circles. Also, by making use of such idioms, some writers may be, to some extent, distancing their work from the general readership and associating it with a more sophisticated, at least language-wise, one.

Selecting Arabic idioms that are loan translations of Western, especially English, ones, could sometimes indicate the speaker's or writer's interest in exhibiting their trendiness and acquaintance with "modernity". Also, particularly in some contemporary literary texts, some writers may employ idioms that are loan translations and, at the same time, evoke historical or religious implications that are shared interculturally. This, in some contexts at least, can be regarded as a way of creating solidarity with even a wider, more inclusive audience.

This function, i.e. creating and reinforcing rapport, is relevant even when language users alter the base form of the expression through idiom-based wordplay (see section 5.1). As Langlotz (2006) notes, this type of wordplay "reasserts the intimate sense of community between the speaker and the hearer since it presupposes the mutually-known neutral base form as a shared starting point for understanding" (p. 210).

Moon (1998) points out that negatively evaluative multiword expressions are "periphrastic devices to convey negative evaluations more politely and less overtly or face-threateningly than simplex items" (p. 265). However, it is important to note that in the case of some negatively evaluative Arabic idioms, even though the evaluative content is conveyed indirectly, this does not mean that it is necessarily expressed politely (see section 5.4.1.1).

Two reasons may be provided for this. First, establishing the form-meaning relationship through the conventionalization of a figurative expression makes it more "direct", or *on record* in Brown and Levinson's (1987) terms, in expressing its non-literal meaning (see the notion of *conventionalized indirectness* in Brown and Levinson 1987). This could also be supported by the fact that many idioms are not used in their literal meanings. Second, the implicatures that are often evoked by the literal readings of many idioms, especially when their figurative images are manipulated in context, may intensify their evaluative content.

It can be said then that although the development of negatively evaluative idioms could be attributed to a need for indirectness, their actual selection in discourse does not always reflect a wish on the part of the speaker or writer to

avoid direct communication of negative judgments. In other words, although the indirectness of idioms may account, even partially, for why they have come into existence, it may not always explain why they are preferred over literal language in some actual contexts.

The foregoing analysis may not only account for why negatively evaluative idioms are more frequent than positively evaluative ones, but it could also offer an additional explanation for the high frequency of the evaluative function in general. The premises of this explanation can be formulated as follows: (a) examples of indirect language are very often used, or even created, to express negative, face-threatening evaluations, and (b) idioms are prototypical instances of indirect language. Then, it could be expected that many idioms would communicate (negative) evaluations.

Many idioms in the data express modalizing content that often manifests in situating an action, quality, etc. near or at one end of a scale between two opposites, usually, but not always, the extreme one. This may be explained by the fact that idioms are not typically used to refer to regular phenomena or moderate qualities. As Carter (2004) points out, the idiomatic meanings normally represent more intense versions of any apparently equal literal statements (p. 132).

In this context, it may be suitable to refer to the findings of some work on the verbal communication of emotions. For example, Gibbs (1994) points out that research on verbal expression of emotional states showed that “[m]ore intense emotions were described using metaphor much more than in descriptions of milder emotions” (p. 126). With regard to idioms that denote fear in particular, Dobrovolskij and Piirainen (2005) note that, “[a]s a whole, idioms tend to express strong emotions, otherwise one would not use idiomatic means of language. Normally, the speaker uses an idiom if he/she feels that standard non-idiomatic linguistic means are not expressive enough” (p. 154).

This typical property of idioms in addition to the other properties discussed above in accounting for the pervasiveness of the informational and evaluative functions in the data can help understand why many idioms perform more than one discursive function simultaneously.

The distribution of information types over the syntactic classes of idioms is in accordance with the general meaning categories usually linked to the syntactic heads of idioms. For instance, actions are typically denoted by verbs, entities by nouns, and qualities by adjectives.

The organizational function is carried out only by prepositional and adverbial idioms. This could be explained by the fact that prepositions and adverbs usually express different types of spatial, temporal, and logical relationships.

Analyzing the figurative image could help see how some idioms contain a modalizing ingredient. For instance, in (36a) “the top of the ice mountain”, the phenomenon is estimated as a small part of a bigger problem by mapping it only onto the tip of the iceberg in the literal scene.

Reworking the figurative image to alter the degree of the evaluative content of the idiom, as in (6b) “the middle jewel of a necklace of identical pearls”, is facilitated by the clarity of the motivation of the idiom. In other cases, e.g. in (7a) “the

place where the mare is tied up”, such a manipulation has not occurred, probably because of the opaqueness of the underlying motivation.

In the foregoing discussion, both isomorphism and motivation have been useful in shedding light on the phenomena related to the discursive functions of Arabic idioms. Yet, the latter seems to be a more powerful explanatory notion in this regard. Isomorphism is mainly about the mappings between the formal constituents of idioms and any parts of their internal semantic structure. Therefore, it should not be surprising to find that this notion does not provide much help with explaining phenomena that are more related to the meaning of the idiom as a whole and how it contributes functionally to its text. The notion of motivation, with its focus on the meaning-meaning relationship rather than the form-meaning one, is a better candidate for this task.

4.2 Idioms and cohesion

Cohesion “can be defined as the set of resources for constructing relations in discourse which transcend grammatical structure” (Martin 2003). “It is the enabling system of ties or links within a text that makes it possible to interpret its elements as meaningful and relevant” (Moon 1998: 278). The relationship between two elements in discourse is often called the cohesive tie. The nature of this tie is semantic, that is, the two elements “are tied together through some meaning relation” (Hasan 1994: 76). As for the relationship between the cohesion of a text and its coherence, Hoey (2001) points out that the “coherence of a text is reflected in and signalled by the cohesion in the text” (p. 51).

Different types of device may be employed to create cohesion in text. These types are discussed in many linguistic accounts of the phenomenon, e.g. Halliday and Hasan (1976), Hasan (1994) and Hoey (1991 and 2001). The list below only provides some examples:

- *simple repetition*: repeating the same lexical item (Hoey 2001);
- *complex repetition*: the use of two morphologically related words such as a verb and a noun that are based on the same stem, e.g. *hibernates* and *hibernation* (Hoey 2001);
- *co-reference*: this subsumes the use of pronouns and determiners;
- *conjunctions*;
- *lexical cohesion*: this refers to the semantic relationships that obtain between the lexis of a text, e.g. synonymy, antonymy and hyponymy.

As far as idioms are concerned, “it is their very lack of cohesion which signals their metaphoricality and anomalousness. The lexico-semantic content of the text sets up contextual constraints whereby any literal value for the metaphor is [normally] excluded” (Moon 1998: 283). This lack of semantic compatibility between the literal reading of the expression and its context is what leads the hearer or reader to resort to the non-literal interpretation in understanding the text (see Croft and Cruse 2004).

The following sections investigate six ways in which Arabic idioms contribute to the cohesion of their texts. Section 4.2.1 examines how the determiners and demonstratives incorporated in some idioms may signal in-text semantic relations. Section 4.2.2 addresses the role of organizational idioms in cohesion. Section 4.2.3 looks at the role of some types of syntactic modifier in creating textual cohesion. Section 4.2.4 explores how extending the figurative images of some idioms contributes to cohesion. Section 4.2.5 looks at the phenomenon of contextual transformations in its relation to cohesion in discourse. Finally, section 4.2.6 addresses how some morphological properties of the idiomatic components may indicate cohesion in text.

4.2.1 Cohesion through determiners and demonstratives

The definite articles incorporated in some idioms may signal information that is given in the co-text or can be retrieved from the context. Since retrieving this information often requires an awareness of longer stretches of the text, some of the examples below are rather long:

- (41) a. *iṣṭāda* *fi* *l-māʾ-i* *l-ʿakir-i*
 fish.PST in DEF-water-GEN DEF-cloudy-GEN
 he fished in troubled waters
- b. *taḍrib-u* *lshawwā* *bi-l-raʿīs-i* *l-sūriyy-i*
 give.as-IND Al-Shawwa with-DEF-president-GEN DEF-Syrian-GEN
ḥāfiẓ *alʿasad* *l-mathal-a* *fi* *tafwīt-i*
 Hafiz Al-Asad DEF-example-ACC in making.escape-GEN
l-furṣat-i *ʿalā ʿisrāʾīl* *bi-l-ṣayd-i* *fi l-māʾ-i*
 DEF-chance-GEN on Israel to-DEF-fishing-GEN in DEF-water-GEN
l-ʿakir-i *ʿindamā dhahaba* *ʿilā ʿammān li-taqdīm-i*
 DEF-cloudy-GEN when go.PST.M.3SG to Amman for-giving-GEN
wājib-i *l-ʿazāʾ-i* *li-l-malik-i* *l-jadīd-i*
 duty-GEN DEF-consolation-GEN to-DEF-king-GEN DEF-new-GEN
wa-l-shaʿb-i *l-ʿurduniyy-i* *l-shaqīq-i*
 and-DEF-people-GEN DEF-Jordanian-GEN DEF-brother-GEN
 (Al-Ahram, 1999)

lit. Al-Shawwa gives the Syrian president Hafiz Al-Asad as an example in making the chance to fish in the cloudy water escape from Israel when he went to Amman to give the duty of consolation to the new king and the brother Jordanian people

Al-Shawwa points to the Syrian president Hafiz Al-Asad, who set an example in depriving Israel of a chance to fish in troubled waters when he went to Amman to console the new king and the sister Jordanian nation

- (42) a. *dassa* *l-summ-a* *fi* *l-ʿasal-i*
 put.secretly.PST DEF-poison-ACC in DEF-honey-GEN
 he deceptively promoted something bad (e.g. an idea or proposal) by presenting it in an attractive way

- b. 'inna 'amaliyyāt-i taḥrīr-i l-mubādālāt-i
indeed activity.F.PL-GEN freeing-GEN DEF-exchange.PL-GEN
l-tijāriyyat-i fi l-sila'-i wa-l-khadamāt-i
DEF-commercial-GEN in DEF-commodity.PL-GEN and-DEF-service.PL-GEN
taṣṭadim-u bi-l-maṣāliḥ-i l-iqtisādiyyat-i
clash.PRS-IND with-DEF-interest.PL-GEN DEF-economic-GEN
li-l-duwal-i l-nāmiyyat-i [. . .] wa-ẓalla
of-DEF-country.PL-GEN DEF-developing-GEN and-remain.PST
l-'ālam-u fi ḥālat-i ghaybūbat-i-n wa-ftiqād-i-n
DEF-world-NOM in state-GEN coma-GEN-INDF and-losing-GEN-INDF
li-l-tawāzun-i khilāla l-tis'iniyyāt-i ḥaythu kāna-t
of-DEF-balance-GEN during DEF-ninety.PL-GEN when be.PST-F
'adawāt-u l-'awlamat-i l-nashīṭat-i taduss.PRS-u
tool.PL-NOM DEF-globalization-GEN DEF-active-GEN secretly.put-IND
l-summ-a fi l-'asal-i wa-tuqaddim-u
DEF-poison-ACC in DEF-honey-GEN and-present.PRS-IND
li-l-'ālam-i ṣūrat-a-n wardiyyat-a-n li-mustaqbal-i
to-DEF-world-GEN picture-ACC-INDF rosy-ACC-INDF of-future-GEN
l-'ālam-i fi ḥill-i l-'awlamat-i (Al-Ahram, 1999)
DEF-world-GEN in shadow-GEN DEF-globalization-GEN
lit. indeed, the activities of freeing trade exchange in products and ser-
vices clash with the economic interests of developing countries [. . .]
And the world remained in a state of unconsciousness and imbalance
during the nineties when the active tools of globalization were secretly
putting the poison in the honey and presenting to the world a rosy picture
of its future in the shadow of globalization
indeed, freeing trade in products and services clashes with the economic
interests of developing countries [. . .] And the world remained oblivious
and unstable during the nineties when the active tools of globalization were
promoting it in an attractive package and presenting to the world a rosy
picture of its future under globalization

In (41b), *al-mā'-u l-'akir* “the cloudy water” refers to some tension in the relationships between Syria and Jordan that Israel, according to the writer, tried to gain an advantage from. In (42b), *al-summ* “the poison” stands for the major economic risks involved in globalization for developing countries, and *al-'asal* “the honey” maps onto the attractive package in which globalization is presented or onto some potential benefits that these countries might gain from it.

The idioms in both examples are isomorphic. The nominal groups in this class of idioms typically bear identifiable parts of the idiomatic meaning, and the entities for which they stand figuratively can be retrieved from the linguistic or extralinguistic context. This explains why the definite articles in these nominal groups can have a signaling function in context.

On the other hand, this type of form-meaning connection is not present in non-isomorphic idioms and, as a result, normally no signaling function can be

identified for the definite articles they contain. Examples (43) to (47) are non-isomorphic idioms in which the definite articles are not used to signal any given or shared information in discourse:

- (43) *tajammada* *l-dam-u* *fi* *‘urūq-i* *shakhs*
 freeze.PST DEF-blood-NOM in vein.PL-GEN *sb*
 the blood froze in *sb*'s veins
- (44) *’adāra* *l-khadd-a* *l-’ākhar-a*
 turn.PST DEF-cheek-ACC DEF-other-ACC
 he turned the other cheek
- (45) *ghuṣn-u* *l-zaytūn-i*
 branch-NOM DEF-olive-GEN
 olive branch
- (46) *marbaṭ-u* *l-faras-i*
 place.where.sth.is.tied.up-NOM DEF-mare-GEN
 the central point (e.g. in a problem or an argument)
- (47) *sābaqa* *l-rīḥ-a*
 race.PST DEF-wind-ACC
 he ran very fast

However, even in the case of isomorphic idioms, the definite article may not signal any shared or given information, e.g.:

- (48) *qimmat-u* *jabal-i* *l-jalīd-i*
 summit-NOM mountain-GEN DEF-ice-GEN
 the tip of the iceberg
- (49) *ḥāraba* *ṭawāḥīn-a* *l-hawā’-i*
 fight.PST mill.PL-ACC DEF-air-GEN
 he tilted at windmills

In (48), the nominal group *al-jalīd* “the ice” does not seem to map onto any part of the idiomatic meaning. Rather, the mapping is between the whole noun phrase *jabal-u l-jalīd* “the iceberg”, lit. the mountain of the ice, and the meaning “a much bigger problem”. In (49), the nominal group *al-hawā’* “the air” does not map onto any part of the idiomatic meaning. Instead, the mapping is between the whole noun phrase *ṭawāḥīn-u l-hawā’* “the windmills”, lit. the air-powered mills, and the meaning part “imaginary enemies”. Therefore, the definite articles in both *al-jalīd* and *al-hawā’* do not perform any signaling function.

In isomorphic idioms, when the reference of the nominal group is already established in context, demonstratives may be used within the idiom structure to signal these referents (Fellbaum 1993), e.g.:

- (50) a. *wāsiṭat-u* *l-’iqd-i*
 middle.part-NOM DEF-necklace-GEN
 the most important/best part of something good

- (53) *fī* *ḍawʿ-i* *shayʿ*
 in light-GEN *sth*
 in the light of *sth*
- (54) *janb-a-n* *ʿilā* *janb-i-n*
 side-ACC-INDF by side-GEN-INDF
 side by side (i.e. “together” in the non-physical sense)

4.2.3 Cohesion through modifiers

Different types of lexical item may be used to modify parts of the idiom. One of the functions that these modifiers can have is to link the idiom to its context (see e.g. Hümmer 2007, on *contextual anchoring*). The modifiers may fill open slots in the structure of the idiom, but sometimes they are not required for the syntactic well-formedness of the expression.

In the following examples, the modifying adjectives and noun phrase contribute to the cohesion of the text by topicalizing the idiom or linking its parts to the context:

- (55) a. *baqarat-u-n* *ḥalūb-u-n*
 cow.F-NOM-INDF milch-NOM-INDF
 milch cow
- b. *wa-baqarat-u-n* *ḍarībiyyat-u-n* *ḥalūb-u-n* (Al-Hayat, 1997)
 and-cow-NOM-INDF related.to.taxes-NOM-INDF milch-NOM-INDF
 lit. and a tax milch cow
- (56) a. *al-yad-u* *l-ʿulyā*
 DEF-hand.F-NOM DEF-upper.F
 the upper hand
- b. *yaqtaḍi* *l-iḥtifāz-a* *bi-l-yad-i*
 require.PRS.M.3SG DEF-keeping-ACC with-DEF-hand.F-GEN
l-ʿamrīkiyyat-i *l-ʿulyā* *ʿalā* *biḥār-i* *l-mantiqat-i*
 DEF-American.F-GEN DEF-upper.F over sea.PL-GEN DEF-region-GEN
 (Al-Hayat, 1997)
 it requires keeping the American upper hand over the seas of the region
- (57) a. *wāsiṭat-u* *l-ʿiqd-i*
 middle.part-NOM DEF-necklace-GEN
 the most important or best part of something good
- b. *fī* *l-kunghris* *lladhī* *yuʿadd-u* *wāsiṭat-a*
 in DEF-Congress REL consider.PRS.PASS-IND middle.part-ACC
ʿiqd-i *l-thawrat-i* *l-jumhūriyyat-i* *llatī*
 necklace-GEN DEF-revolution-GEN DEF-republican-GEN REL
sayṭara-t *ʿala* *l-wilāyāt-i*
 take.control.PST-F on DEF-state.PL-GEN
l-muttaḥidat-i (Al-Thawra, after 2000)
 DEF-united-GEN
 lit. in the Congress that is considered the middle jewel of the necklace of the republican revolution that took control of the United States

In (55b), the adjective *ḍarībiyyah* “related to taxes” links the idiom to the topic of the text. In (56b), the adjective *ʿamrīkiyyah* “American” states the entity in control and links the idiom to context. In (57b), the noun phrase *al-thawrat-u l-jumhūriyyat-u llatī sayṭara-t ʿala l-wilāyāt-i l-muttaḥidah* “the republican revolution that took control of the United States” is a genitive modifier that states the discourse entity for which the noun *ʿiqd* “necklace” stands metaphorically.

The idioms in these cases are isomorphic and can therefore be modified in the way specified in order to help create cohesion. However, the issue of syntactic modification in idioms is a very complicated one that is worthy of more in-depth investigations. Indeed, syntactic modifiers may contribute to cohesion, among other functions, in different ways and with idioms of different internal semantic structures (for some recent accounts of the phenomenon in other languages see Stathi 2007 on German verbal idioms and Minugh 2007 on English idioms).

4.2.4 Cohesion through extending the figurative image

Sometimes, the image underlying the idiom is developed in the co-text by using lexical items that are not part of its base form and that belong semantically to the source domain of the metaphor (see Moon 1998). Through using these items, the speaker or writer creates cohesive ties at the literal level between the idiom and its co-text.

Extending metaphors can be motivated by a wish to create humor or to be sarcastic, as in (58b), but this is not always the case, e.g. in (59b) and (60b):

- (58) a. ʿadāra l-khadd-a l-ʾākhar-a
 turn.PST DEF-cheek-ACC DEF-other-ACC
 he turned the other cheek
- b. ʾidhā ʾadarnā la-hu khadd-a-na l-ʾayman-a
 COND turn.PST 1PL to-M.3SG cheek-ACC-1PL.POSS DEF-right-ACC
 baʿda ʾan yaḍrib-a khadd-a-na l-ʾaysar-a
 after COMP hit.PRS.M.3SG-SBJV cheek-ACC-1PL.POSS DEF-left-ACC
 fa-sawfa yuballiṭ-u-nā ʿalā muʾakhirat-i-nā
 so-FUT bring.down.PRS.M.3SG-IND-1PL on bottom-GEN-1PL.POSS
 (Al-Thawra, after 2000)
 lit. if we turn to him our right cheek after he hits our left one, he would bring us down on our bottoms
- (59) a. qimmat-u jabal-i l-jalīd-i
 top-NOM mountain-GEN DEF-ice-GEN
 the tip of the iceberg
- b. yumkin-u rasm-u baʿd-i
 be.possible.PRS-IND drawing-NOM some-GEN
 taḍāriṣ-i qimmat-i jabal-i l-jalīd-i
 topographical.feature.PL-GEN summit-GEN mountain-GEN DEF-ice-GEN
 l-ṭāfiyat-i ʿala l-saṭḥ-i (Al-Hayat, 1997)
 DEF-floating-GEN on DEF-surface-GEN

lit. it is possible to draw some of the topographical features of the top of the ice mountain that is floating on the surface

- (60) a. 'ibrat-u-n fī kawmat-i qashsh-i-n
 needle-NOM-INDF in stack-GEN hay-GEN-INDF
 a needle in a haystack
- b. 'inna l-qitāl-a l-fa'āl-a ḍidda l-'irhāb-i
 indeed DEF-fighting-ACC DEF-effective-ACC against DEF-terrorism-GEN
 laysa l-baḥṭh-a 'an-i l-'ibrat-i fī kawmat-i
 NEG DEF-searching-ACC for DEF-needle-GEN in stack-GEN
 l-qashsh-i bal 'izālat-a kawmat-i l-qashsh-i kull-i-hā
 DEF-hay-GEN but removing-ACC stack-GEN DEF-hay-GEN all-GEN-F.3SG
 (Al-Watan, 2002)

lit. indeed, effective fighting against terrorism is not searching for the needle in the haystack, but it is removing the haystack altogether

In (58b), *'ayman* “right”, *yaḍrib* “to hit”, *'aysar* “left”, and *yuballīṭ-u-nā 'alā mu'akhirat-i-nā* “he brings us down on our bottom” are semantically compatible with the literal reading of the idiom. In particular, the latter intensifies the sense of humiliation usually expressed by the idiom and may be a source of humor for some readers, i.e. a sort of “black humor”. In (59b), *rasm-u ba'ḍ-i taḍārīs* “drawing some topographical features” and *al-ṭāfiyat-u 'ala l-saṭḥ* “floating on the surface” relate semantically to the literal reading of the expression. A comment along the same lines can be made about *'izālah* “removing” in (60b).

Such extensions sometimes intensify the idiomatic meaning, e.g.:

- (61) a. shadda l-ḥizām-a
 tighten.PST DEF-belt-ACC
 he tightened his belt
- b. al-muwāṭin-ūn yashudd-ū-na l-'aḥzimat-a
 DEF-citizen.PL-NOM tighten.PRS-3PL-IND DEF-belt.F.PL-ACC
 'alā 'ākhir-i-hā (Al-Watan, 2002)
 to end-GEN-F.3PL.POSS
 lit. citizens are tightening their belts to the maximum

In (61b), *'alā 'ākhir-i-hā* “to their end point” expresses the great extent to which citizens reduce their expenditure.

Generally, in order for the language user to be able to deploy such extensions, the motivation of the idiom needs to be clear to them. The effects that these extensions have may be attributed to the vivid images and implicatures of the elements added. Conventionalization could reduce the vividness of the figurative motivation. Extending the figurative images by using non-conventional elements, on the other hand, brings them back to life and allows the language user to add to or make changes in the idiomatic meanings. Langlotz (2006) notes that “[b]y re-constructing the literal scene by means of conspicuous, non-core vocabulary, its conceptual and emotional entailments can be further exploited

to trigger a series of additional weakly implicated premises and conclusions” (p. 210).

4.2.5 Cohesion through contextual transformations

Occasionally, some idiomatic components may be replaced with lexical items that refer literally or even metaphorically to what these components represent in discourse. These transformations make the idiom cohesive with its co-text by showing how the idiomatic parts are interpreted in context.

Generally, finding examples of this phenomenon in *Arabic corpus* has been a matter of chance, as it is not possible to predict in advance whether, and in what way, a certain idiom would be used in this manner, e.g.:

- (62) a. *ḍaraba* *‘uṣfūr-ay-ni* *bi-ḥajar-i-n*
hit.PST sparrow-DU.ACC-INDF with-stone-GEN-INDF
he killed two birds with one stone
- b. *sawfa* *yuṣīb-u* *‘uṣfūr-ay-ni* *bi-qānūn-i-n*
FUT hit.a.target-IND sparrow-DU.ACC-INDF with-law-GEN-INDF
wāḥid-i-n (Al-Ahram, 1999)
one-GEN-INDF
lit. it will hit two sparrows with one law
- c. *ya‘nī* *takrīs-a* *ḥaqīqat-i-n* *‘isrā’īliyy-at-i-n*
mean.PRS.M.3SG establishing-ACC fact-GEN-INDF Israeli-F-GEN-INDF
jadīd-at-i-n *‘ukhrā* *wa-‘izāḥat-a* *ḥaqā’iq-a*
new-F-GEN-INDF another and-removing-ACC fact.PL-GEN
filistīniyy-at-i-n *bi-ḥajar-i-n* *wāḥid-i-n*
Palestinian-F-GEN-INDF with-stone-GEN-INDF one-GEN-INDF
(Al-Hayat, 1997)
lit. it means establishing another new Israeli fact and removing
Palestinian facts with one stone

In (62b) and (62c), contextual transformations have been applied to the object of the preposition and to the verb and its object, respectively. In another idiom:

- (63) a. *‘ibrat-u-n* *fī* *kawmat-i* *qashsh-i-n*
needle-NOM-INDF in stack-GEN hay-GEN-INDF
a needle in a haystack
- b. *yarsim-u* *l-shā‘ir-u* *khuṭūt-a-n* *li-malāmiḥ-i*
draw.PRS-IND DEF-poet-NOM line.PL-ACC-INDF of-feature.PL-GEN
l-gharīb-i *l-hā’im-i* *l-bāḥith-i* *‘an*
DEF-stranger-GEN DEF-wanderer-GEN DEF-searcher-GEN for
wajh-i-hi *fī* *kawmat-i* *qashsh-i-n* *gharībat-i-n*
face-GEN-M.3SG.POSS in stack-GEN hay-GEN-INDF strange-GEN-INDF
‘an *rūḥ-i-hi* (Al-Hayat, 1996)
to soul-GEN-M.3SG.POSS

l-rūs-i	lladhīna	fāz-ū	‘alayhim	fi
DEF-Russian.PL-GEN	REL	achieve.victory.PST-3PL	upon.M.3PL	in
nihā’iyy-i	buṭūlat-i	l-‘ālam-i	l-māḍiyat-i	bi-lyābān
final-GEN	championship-GEN	DEF-world-GEN	DEF-last-GEN	in-Japan
wa-fi	nafs-i	l-waqt-i	jaddad-ū	’aḥlām-a-humu
and-at	same-GEN	DEF-time-GEN	renew.PST-3PL	dream.PL-ACC-M.3PL.POSS
l-qadīmat-a	bi-’an	yakūn-ū	’asyād-a	l-‘ālam-i
DEF-old-ACC	of-COMP	be.PRS-3PL.SBJV	master.PL-ACC	DEF-world-GEN
fi	hādhihi	l-lu‘bat-i		
in	DEM	DEF-game-GEN		

the Swedish snatched the sixteenth handball world championship and killed three birds with one stone: they won the title for the fourth time in their life, they took revenge on the Russians who defeated them in the final at the last world championship in Japan, and, at the same time, they revived their old dreams of being the masters of the world in this game

In this example, *hajar-u-n wāḥid* “one stone” stands metaphorically for the Swedish victory in the final match at the handball world championship, while *thalāthat-u ‘aṣāfir* “three sparrows” maps on winning the title for the fourth time, taking revenge on the Russian team, and reviving the dreams of being the best in this game.

Another idiom that can be used as a good example here too is:

- (66) bayna nār-ayni
 between fire-DU.GEN.INDF
 between the devil and the deep blue sea

In its corpus examples, the two difficult choices, referred to metaphorically by *nārān* “two fires”, are normally stated clearly in the co-text. The isomorphism of the two examples evidently accounts for the possibility to notice these relationships between their formal characteristics and relevant elements in their contexts.

Once more, both isomorphism and motivation have been useful in explaining the phenomena related to the role of idioms in creating cohesion. However, unlike the case with the discursive functions of idioms, isomorphism seems to be a more powerful explanatory notion in this context. As noted earlier, cohesion is realized through formal features in text. A cohesive tie links one formal element to another in a text. Due to its focus on the connections between the formal structure of the idiom and its semantic structure, isomorphism is generally a better candidate for accounting for the occurrence of many cohesive ties between the idiomatic components and other formal elements in context.

5 The lexical and grammatical behavior of Arabic idioms

This chapter addresses the formal properties of idioms as manifested in their lexical and grammatical behavior. It begins by introducing a basic distinction between the notions of *systematic variation* and *wordplay*, relying on some recent works in the linguistic literature on idioms.

Since some aspects of the formal behavior of idioms cannot be deemed purely lexical or purely grammatical, this chapter deals with the relevant phenomena in the data within three major sections: lexical behavior of idioms, lexicogrammatical behavior, and grammatical behavior. In the first, the point at issue is lexical variation in idioms. The second section is concerned with two main issues: changes in lexicogrammatical complexity, and perspective-adaptation.

The third section focuses on various aspects of the grammatical behavior of the data. It is broken down into three major parts. The first pays attention to the morphological behavior of idioms as manifested mainly in their potential for inflection. The second addresses itself to the use of the active and passive voice. Finally, the third part deals with the syntactic behavior of the data with regard to their syntactic types and structures, constituent order, embedding, and predicatization.

In the discussions to follow, I have attempted to link (any in-text changes in) the formal structures of idioms to their semantic and discursive properties as set out in Chapters 3 and 4, and, when appropriate, to the semantic, grammatical, and discursive properties of their texts.

5.1 Systematic variation vs. wordplay

Before going into the details of the various aspects of the lexical and grammatical behavior of the data, it is important to distinguish between these two types of formal variation in idioms. This distinction is essential for proposing any general explanations or drawing up any rules that govern the in-text formal behavior of idioms. In what follows, I mainly rely on Langlotz (2006) for his detailed account of the phenomenon.

Systematic formal variation in idioms refers to changes that are limited by the idiom's formal and semantic structure. These changes "do not trigger strikingly unconventional semantic effects" (Langlotz 2006: 223). They mainly aim to

maximize the semantic integration of the idiom into its context and determine the specific meaning that it takes on in a particular text.

Therefore, these variations could be explained by general rules that are derived mainly from the properties of the semantic structure of idioms and from their discursive characteristics. Such rules may help understand why certain types of variation occur with certain idioms but not with others and enable to appreciate the semantic and/or discursive effects that might come about as a result of such variations.

Unlike systematic variation, wordplay often triggers semantic effects that are strikingly different from the conventional idiomatic meaning. It is not limited by the formal and semantic structure of the idiom and, therefore, is not predictable and can occur practically with any idiom (Langlotz 2006). In Langlotz's words:

[w]ith wordplay, the formal and semantic structure of an idiom is exploited as a resource for stylistic and poetic effects that also communicate the speaker's intention to engage in humour, to be conspicuous and trigger additional, weakly implicated facets of meaning. (p. 201)

He discusses two major types of wordplay in idioms:*

Idiom variation must be regarded as wordplay if a given idiom alternate is used to trigger a series of weakly-implicated semantic or stylistic effects that go beyond the systematic contextual integration of the idiomatic meaning in discourse. This is the case when the idiom is intentionally ambiguated or changed so that it becomes stylistically marked and inconsistent with its conventional idiomatic form and meaning. Wordplay can also be effected without affecting the integrity of the idiomatic meaning through parasitic elaboration. This is the case when a given pun variant cannot activate the idiomatic meaning independently, but depends on the previous recognition of the idiom in accordance with its conventional usage. (p. 203)

In relation to the former type, he notes that:

semantic variation through ambiguation is not conventionalised, but an occasional meaning adaptation. [. . .] By the process of ambiguation both levels of meaning – the idiomatic meaning and a potential literal meaning – have their own reference in the usage-context. (p. 181)

The following advertising slogan illustrates this type:

(1) If you like wine, have a good nose.

The idiom *have a good nose for sth.* ("be good at discovering precious and interesting things") is ambiguated by activating the literal meaning of *have a good nose*. A wine expert, who is good at discovering interesting and high-quality products, must also depend on their nose.

(Langlotz 2006: 181)

Langlotz illustrates the notion of *parasitic elaboration* by the following example, taken from Glucksberg (2001):

(2) Speaker A: Did the old man kick the bucket last night?

Speaker B: Nah, he barely nudged it.

In the reply, Speaker B plays with the image evoked by the literal meaning of *kick the bucket*. Instead of negating the idiomatic meaning (“no, he didn’t die”), the negation is effected by manipulating the literal scene in terms of an inferential analogy: *kick the bucket: not nudge the bucket = die: not die*. This analogical modification is not prevented by the specific nature of the idiom’s formal and semantic structure – *kick the bucket* is a notoriously opaque idiom – but rather depends on the accessibility of the literal scene in discourse.

(Langlotz 2006: 202)

Having set this distinction, it should be noted that the boundaries between systematic variation and wordplay in idioms might sometimes become blurred. This could occur, for example, when different speakers motivate the same idiom in different ways, as pointed out in Chapter 3 (and see Langlotz, 2006: 205).

It should be stated here that the focus in this work is on systematic variation, since this type is explainable by the general and specific semantic and discursive qualities of the data.

5.2 Lexical variation

Many idioms in the data allow for substituting one or more of their base form lexis with other literal or non-literal lexical items. In more technical parlance, the corpus uses of these idioms violate some of the usual paradigmatic constraints on the selection of their lexical elements. In such variations, the essence of the idiomatic meaning is preserved. However, the variants might differ with regard to their degree of formality or the text types they typically occur in. Although this might be an interesting point to examine in detail, the focus here is on how idioms vary lexically and how this could be explained mainly by their semantic properties.

This section is only concerned with simple substitutions of the lexical components of idioms. Lexical changes that come about as part of more complex lexicogrammatical operations or in the course of some in-text reworkings of the figurative images are discussed in other appropriate sections (see e.g. section 5.3.2).

It has become evident from the analysis that some substitutions occur more often than others do. Therefore, these seem to constitute alternative institutionalized variants of their respective idioms. On the other hand, other substitutions occur only once or twice in the corpus. No decision about institutionalization can be made on the basis of this low frequency alone. This also reminds of Fellbaum’s (2007b) remark that lexical variation in idioms “is usually discourse-conditioned and ad-hoc” (p. 12). In a relevant context, Langlotz (2006) terms the former type of variants *usual variants*, contrasting them with *occasional variants*.

Although the numbers given in the following sections report both types of lexical variation, i.e. the usual and occasional, the examples themselves often represent the cases where the substitutions seem to have developed some degree of institutionalization in the speech community, as indicated mainly by their relatively high frequency in the data.

Sections 5.2.1 to 5.2.4 provide instances of lexical variation in the verbal, nominal, adjectival, and prepositional components of idioms, respectively. For convenience, these instances are discussed together in section 5.2.5. In the examples below, some of the alternative lexical items are presented and separated by forward slashes. In presenting these items, I have usually started with the ones that occur in the base forms of the idioms as given in Chapter 2.

5.2.1 Verb variation

Seventeen idioms in the data exhibit verb variation: three verb-subject idioms and fourteen verbal ones. Broadly synonymous variants occur through replacing the verb with another (synonymous) non-literal verb that may or may not be derived from the same root, as in examples (3) to (10), or through replacing the verb with another that conveys literally its share of the idiomatic meaning, as in (11). Also, in some polysemous idioms, some variations appear to be only possible with one of the idiomatic meanings but not with the other(s). This applies, for example, to the use of *jaffa* “to dry” in (12b), but not in (12a):

- (3) *shay*'_i / 'akhadha / 'amsaka ba'ḍ-u-*hu*_i bi-riqāb-i
*sth*_i hold.on.PST / hold.PST some.of-NOM-M.3SG.POSS_i to-neck.PL-GEN
 ba'ḍ
 some
*sth*_i (e.g. a set of ideas, texts, or problems) interconnected
- (4) *dassa* / waḍa'a l-summ-a fi l-'asal-i
 put.secretly.PST / put.PST DEF-poison-ACC in DEF-honey-GEN
 he deceptively promoted something bad (e.g. a way of thinking or behaving)
 by presenting it in an advantageous or attractive way
- (5) ḥāraba / ṣāra'a ṭawāḥīn-a l-hawā'_i
 fight.PST / wrestle.with.PST mill.PL-ACC DEF-air-GEN
 he tilted at windmills
- (6) 'adāra / 'a'ṭa l-khadd-a l-'ākhar-a
 turn.PST / give.PST DEF-cheek-ACC DEF-other-ACC
 he turned the other cheek
- (7) iṣṭāda / ṣāda fi l-mā'_i l-'akir-i
 fish.PST / fish.PST in DEF-water-GEN DEF-cloudy-GEN
 he fished in troubled waters
- (8) 'amsaka / masaka l-'aṣā min-a l-muntaṣaf-i
 hold.PST / hold.PST DEF-stick from DEF-middle-GEN
 he struck a balance

- (9) bala‘a / ibtala‘a l-ṭu‘m-a
 swallow.PST / swallow.PST DEF-bait-ACC
 he swallowed the bait
- (10) ḍaraba / ‘aṣāba / iṣṭāda / ṣāda ‘uṣfūr-ayni
 hit.PST / hit.a.target.PST / catch.PST / catch.PST sparrow-DU.ACC.INDF
 bi-ḥajar-i-n
 with-stone-GEN-INDF
 he killed two birds with one stone
- (11) dhāqa / ‘ānā / qāsā marārat-a *shay’*
 taste.PST / suffer-PST / suffer-PST DEF-bitterness-ACC *sth*
 he experienced the suffering of *sth*
- (12) a. tajammada / jamada l-dam-u fī ‘urūq-i *shakḥ*
 freeze.PST / freeze.PST DEF-blood-NOM in vein.PL-GEN *sb*
 the blood froze in *sb*'s veins
 b. tajammada / jamada / jaffa l-dam-u
 solidify.PST / solidify.PST / dry.PST DEF-blood-NOM
 fī ‘urūq-i *shay’*
 in vein.PL-GEN *sth*
 lit. the blood solidified or dried in *sth*'s veins
sth (e.g. an organization or political movement) became inactive

5.2.2 Noun variation

Twelve idioms show noun variation. In a similar way to verb variation, in some cases, the base-form noun and the substitute are derived from the same root, e.g. in (13) and (14), in other cases, e.g. (15) to (18), they are not:

- (13) *shay’* marra murūr-a / marr-a l-kirām-i
sth pass.PST passing-ACC / passing-ACC DEF-noble.people-GEN
 lit. *sth* (e.g. an event or anniversary) passed as noble people do
sth passed unnoticed
- (14) dhāqa marārat-a / murr-a *shay’*
 taste.PST bitterness-ACC / bitterness-ACC *sth*
 he experienced the suffering of *sth*
- (15) qarṣat-u / shaddat-u ‘udhun-i-n
 pinch-NOM / pull-NOM ear-GEN-INDF
 a light non-physical punishment that serves as a warning of a possible
 cruel one
- (16) ‘affara jabhat-a-hu / jabīn-a-hu
 cover.with.dust.PST forehead-ACC-M.3SG.POSS / forehead-ACC-M.3SG.POSS
 he humiliated himself
- (17) qimmat-u jabal-i l-jalīd-i / l-thalj-i
 top-NOM mountain-GEN DEF-ice-GEN / DEF-ice-GEN
 the tip of the iceberg

5.2.5 Discussion of lexical variation

The transparency of the underlying motivation of the idioms and of their internal semantic make-up may account for most of the lexical substitutions reported above. In her analysis of the phenomenon within the Head-Driven Phrase Structure Grammar framework, Riehemann (2001) notes that:

[i]n the case of idioms based on active metaphors it is [. . .] desirable to express the more general metaphors, in order to be able to deal with productive substitutions of words in these idioms. For example if the more general metaphor underlying the idiom *break the mold* is expressed as a higher-level constraint, then productive uses like *shatter the mold*, *crack the mold*, and *break out of the mold* can be understood based on the same metaphor.

(pp. 37–38, and see Gibbs 1994: 283ff)

While replacing the idiomatic word with another non-literal lexical item can be sufficiently explained by the transparency of the motivation of the idiom, e.g. in (16) and (19), replacing it with a lexical item that literally conveys part of the idiomatic signification normally entails that the idiom is isomorphic and that the idiomatic word in question bears a discernible share of its meaning, as in (11).

As noted in section 5.2.1, the verb *jaffa* “to dry” is used with the idiom in (12) in only one of its two meanings, i.e. in (12b) but not in (12a). The two meanings appear to be based on two different processes of semantic extension. As discussed in Chapter 3, the idiomatic reading “to be frightened” in (12a) is mainly based on a metonymic connection that manifests in an (assumed) drop in the temperature of the blood when people feel afraid. On the other hand, the meaning of “inactivity” in (12b), e.g. of an organization or a treaty, stems from a perceived similarity between an inactive organization, political party, etc. and a dead body that shows no blood circulation. Given this analysis, the literal meaning of *jaffa* can only be relevant to the figuration underlying the latter meaning, which could explain why it has occurred only with the idiom when used in this meaning.

On the other hand, as Fellbaum (2007b) points out, “lexical selection is even stronger in expressions that are not semantically transparent” (p. 9). However, in (18), noun variation occurs in an idiom that is not motivated for many native speakers. The variant with the substitute has a literal meaning which is very similar to that of the base form. This variation could be explained by the fact that the idiom is literally compositional.

Finally, in (27), the variant that is almost always cited in Arabic dictionaries is the one headed by *‘alā* “on”. This is by far the less frequent variant of the idiom in the corpus. Evidently, the meaning of *‘alā* is relevant to the original context of the literal reading of the expression (see section 3.2.6). Both variants are literally non-compositional mainly due to the presence of *bakrah*, however (see section 3.2). It seems that this absence of any knowledge of the literal meaning of the expression and the lack of any obvious relationship between the meaning of *‘alā* and the idiomatic meaning have allowed for the variation in the preposition to

occur. The phonological similarities between the two prepositions could partially account for the use of *'an* “about” rather than any other preposition in this position. Having said that, some diachronic investigations are needed so as to find the answer to how and why this replacement came about in the first place.

5.3 Lexicogrammatical behavior of idioms

5.3.1 Changes in the lexicogrammatical complexity

Five idioms in the data have two formal variants that differ mainly with respect to their lexicogrammatical complexity. For instance, in (28), the prepositional phrase *bi-yad-i-hi* “with his own hand” is sometimes dropped. In (29), the noun *'ab* “father” is sometimes dropped, and, as a result, the possessive pronoun that the context provides becomes a dependent of the object of the preposition, i.e. *bakrah* “young female camel”. The fuller version of this idiom occurs 76 times in *Arabi-corpus*, while the shorter version occurs only twice, showing the same range of preposition variation that the fuller version has:

- (28) a. *ḥafara qabr-a shay' bi-yad-i-hi*
 dig.PST grave-ACC *sth* with-hand-GEN-M.3SG.POSS
 he put an end to *sth*
- b. *'dhā kāna-t-i l-shuyū'īyyat-u l-rūsiyyat-u qad*
 if be.PST-F DEF-communism-NOM DEF-Russian-NOM already
ḥafara-t qabr-a-hā bi-yad-i-hā (Al-Ahram, 1999)
 dig.PST-F grave-ACC-F.3SG.POSS with-hand-GEN-F.3SG.POSS
 if the Russian communism had already dug its grave with its own hand
- c. *ḥadhdhara-t ṭahrān-u min 'anna qādat-a l-jaysh-i*
 warn.PST-F Tehran-NOM of COMP leader.PL-ACC DEF-army-GEN
l-turkiyy-i yaḥfur-ū-na qubūr-a-hum (Al-Hayat, 1997)
 DEF-Turkish-GEN dig.PRS-M.3PL-IND grave.PL-ACC-M.3PL.POSS
 Tehran warned that the leaders of the Turkish army are digging their own graves
- (29) a. *shakḥš/shay' 'alā / 'an bakrat-i 'ab-ī-hi*
sb/sth_i on / about young.camel.F-GEN father-GEN-M.3SG.POSS_i
sb_i (e.g. a group of people) all together or all *sth_i* (also can be used adverbially with some adjectives, participles, and verbs to mean *completely* or *fully*)
- b. *yakhruj-u l-muwāṭin-ūna 'an bakrat-i*
 go.out.PRS-IND DEF-citizen-PL.NOM about young.camel.F-GEN
'ab-ī-him (Al-Ahram, 1999)
 father-GEN-M.3PL.POSS
 lit. the citizens go out about their father's young female camel
 the citizens go out all together
- c. *'aḥzimāt-u ṭalaqāt-i l-banādiq-i [. . .] maḥshuwwat-a-n*
 belt.PL-NOM bullet.PL-GEN DEF-rifle.PL-GEN stuffed-ACC-INDF

‘alā bakarat-i-hā (Al-Hayat, 1996)

on young.camel.F-GEN-F.3SG.POSS
 lit. the belts of the rifle bullets are stuffed on their young female camel
 the belts of the rifle bullets are completely stuffed

Retaining the prepositional phrase in (28b) emphasizes the meaning of self-destruction, since the subject and the possessive that fills the structural slot after the object are co-referential. However, the prepositional phrase could be dropped even if these two elements are co-referential, as in (28c).

In the literal meaning of (28a), the prepositional phrase does not normally add much semantic content. That is, since people usually use their hands to dig or to handle the digging tools, the meaning of the prepositional phrase would normally be inferred even if it is not referred to explicitly. Thus, this phrase is mainly used so as to emphasize the involvement of the particular agent, especially since the context often supplies a possessive that is co-referential with the subject. This may account for why the shortened version does not express the same degree of emphasis found in the fuller one. The clarity of the underlying motivation and the isomorphism of the idiom may account for why the prepositional phrase can be dropped and why dropping it has the particular effect it has.

It is worth mentioning that *yaḥfur-ū-na qubūr-a-hum* “they are digging their own graves” in (28c) is incorporated within a pair of double angle brackets in its original text. This could be explained by one or both of the following. First, these brackets might have been used to mark a quotation, i.e. to indicate that the expression is a literal translation of its source. Second, the context shows that the possibility of a war was not remote, so the brackets may indicate that the expression does not only refer to putting an end to the careers or power of the leaders of the army, but also to their lives, i.e. this use of the brackets might be taken as a way of making the literal meaning of the idiom relevant to the context. However, it should be noted that neither of these seems to be a standard use of this punctuation mark in MSA.

In (29), dropping the second noun in the construct phrase, i.e. *’ab* “father”, does not appear to bring about any semantic or discursive changes in the idiom. The opaqueness of the expression and its very general meaning may account for this.

The other three idioms that show variation in their lexicogrammatical complexity are:

- (30) *shay’* marra murūr-a l-kirām-i bi-*shakhṣ*
sth pass.PST passing-ACC DEF-noble.people-GEN by-*sb*
sth (e.g. an event) passed unnoticed by *sb*
- (31) qalaba zahr-a l-mijann-i li-*shakhṣ/shay’*
 turn.around.PST back-ACC DEF-shield-GEN to-*sb/sth*
 he became hostile to *sb/sth* after he was a friend or supporter of them/it
- (32) min warā’-i zahr-i *shakhṣ*
 from behind-GEN back-GEN *sb*
 behind *sb*’s back

In (30), the prepositional phrase headed by *bi-* “by”, or sometimes ‘*alā* “by”, can be dropped. The pattern of complementation associated with *marra* “to pass” consists of a subject and an optional prepositional phrase that is headed by *bi-* or ‘*alā*. This may account for why this phrase can be left unmentioned.

Functionally, dropping this phrase may be useful when the writer does not want to specify the person or group of people that did not notice the event in question. This may occur simply because specifying these people may not be important in context. Also, since the idiom could express a negative evaluation of that person or group of people, the writer might drop the prepositional phrase to conceal their identity.

In (31), the prepositional phrase headed by *li-* “to” has been dropped sometimes in the examples of this idiom in *Arabicorpus*. Again, the optionality of this phrase in the complementation pattern of the head verb could account for this. Finally, in (32), *min* “from” is sometimes dropped from the structure of the idiom. In MSA in general, *warā* “behind” can be used with or without *min* (see Badawi *et al.* 2004: 218ff).

5.3.2 Perspective-adaptation

In some cases, an idiom may have two or more variants that differ in terms of, for instance, transitivity and intransitivity, causativity, and reflexivity (see Stathi 2007: 107). Langlotz (2006) notes that “a number of verbal idioms vary their lexical or syntactic structure systematically to focus on the scene or action described by the literal meaning from an alternative perspective” (p. 182). He illustrates this by the English expressions *the curtain comes down on sth* and *to bring the curtain down on sth*, where the former is used inchoatively and the latter causatively.

Many of such variations are discussed under the label *systematic variation* in Moon (1998), since they are predictable in many cases in her data. As she puts it, these variations “may be predicted to occur in text, although this does not necessarily mean that they do occur” (p. 90). Her term has not been adopted here, because it is used for a different concept that is introduced in the beginning of this chapter.

Six cases of perspective-adaptation have been identified in the data. All these involve the use of a causative verb in one variant, as in the first of each pair of examples below, and a verb that is reflexive or expresses a spontaneous development (see Ryding 2005: 669ff) in the other, as in the second of each pair of examples below:

- | | | | | |
|---------|---|-------------------|-----------|--------------|
| (33) a. | <i>shay</i> ' | qabaḍa | ṣadr-a | <i>shakḥ</i> |
| | <i>sth</i> | contract.CAUS.PST | chest-ACC | <i>sb</i> |
| | lit. <i>sth</i> made <i>sb</i> 's chest contract | | | |
| | <i>sth</i> (e.g. some news) made <i>sb</i> feel depressed | | | |
| b. | inqabaḍa | | ṣadr-u | <i>shakḥ</i> |
| | contract.REFL.PST | | chest-NOM | <i>sb</i> |

- lit. *sb*'s chest contracted
sb felt depressed
- (34) a. *kasara* shawkat-a *shakhṣ/shay'*
break.CAUS.PST thorn-ACC *sb/sth*
lit. he broke *sb/st*'s thorn
he deprived *sb/sth* of their/its power
- b. *inkasara-t* shawkat-u *shakhṣ/shay'*
break.REFL.PST-F thorn-ACC *sb/sth*
lit. *sb/sth*'s thorn got broken
the power of *sb/sth* was lost
- (35) a. *rafa'a* rāyat-a *shay'*
raise.PST flag-ACC *sth*
lit. he raised the flag of *sth*
he made *sth* triumphant
- b. *irtafa'a-t* rāyat-u *shay'*
go.up.PST-F flag-NOM *sth*
lit. the flag of *sth* went up
sth became triumphant
- (36) a. *jammada* l-dam-a fī 'urūq-i *shakhṣ*
freeze.CAUS.PST DEF-blood-ACC in vein.PL-GEN *sb*
he froze the blood in *sb*'s veins
- b. *tajammada* l-dam-u fī 'urūq-i *shakhṣ*
freeze.PST DEF-blood-NOM in vein.PL-GEN *sb*
the blood froze in *sb*'s veins
- (37) a. 'adāra raḥā *shay'*
spin.CAUS.PST quern.F *sth*
lit. he made the quern of *sth* spin
he made *sth* (e.g. a war or elections) take place in a violent way
- b. *dāra-t* raḥā *shay'*
spin.PST-F quern.F *sth*
lit. the quern of *sth* spun
sth (e.g. a war or elections) took place in a violent way
- (38) a. 'adhāqa-hu marārat-a *shay'*
taste.CAUS.PST-M.3SG bitterness-ACC *sth*
lit. he made him taste the bitterness of *sth*
he made him suffer greatly from *sth*
- b. *dhāqa* marārat-a *shay'*
taste.PST bitterness-ACC *sth*
lit. he tasted the bitterness of *sth*
he experienced the suffering of *sth*

All of these uses may be accounted for by the clarity of the motivation of the idioms. In addition, the meaning of causation could sometimes be expressed phrasally, rather than derivationally as above, through the use of causative verbs such as *ja'ala* "to make". Since this use brings about changes in the syntactic

structure of the idiom by redistributing its components over two clauses – a matrix clause and a subordinate one – its examples have been discussed under *Embedding* in section 5.4.3.3.

5.4 The grammatical behavior of idioms

Free lexical combinations are restricted in their morphological and syntactic behavior mainly by the general grammatical rules of the language. Idioms, on the other hand, could have, in addition to these rules or even sometimes in disagreement with them, other restrictions that stem from their semantic, lexical, discursive, and/or etymological properties.

In this context, Gross' (1993) concept of *local grammars*, i.e. grammatical rules that apply only to a specific expression or group of expressions, is of special relevance. In fact, developing such grammars has been the focus of some previous work, in particular within the domain of natural language processing (see e.g. Breidt *et al.* 1996). The grammatical behavior of idioms is of particular relevance also to construction grammar. As Croft and Cruse (2004) note, “[i]t is not an exaggeration to say that construction grammar grew out of a concern to find a place for idiomatic expressions in the speaker’s knowledge of a grammar of their language” (p. 225).

This section is concerned with investigating the grammatical behavior of Arabic idioms. However, in some cases, it gives attention to the properties of the contexts of these idioms when this is needed for a fuller understanding of the grammatical behavior of the data themselves.

5.4.1 Inflectability of idioms

5.4.1.1 Inflectability of verbs

The verbs in verb-subject idioms inflect fully for tense and follow the general rules of the language concerning their agreement with their subjects. Some corpus examples of these include:

- (39) a. 'afala najm-u *shakhš/shay'*
 set.PST star-NOM *sb/sth*
 the glory or fame of *sb/sth* ended
- b. fa-'inna najm-a-hu sa-ya'ful-u
 and.so-indeed star-ACC-M.3SG.POSS FUT-set-IND
 ka-faylasūf-i-n (Al-Hayat, 1997)
 as-philosopher-GEN-INDF
 lit. and so his star as a philosopher will set
- (40) a. inqabaḍa ṣadr-u *shakhš*
 contract.PST chest-NOM *sb*
 sb felt depressed
- b. wa-tanqabiḍ-u ṣudūr-u-hum 'izā'a kull-i
 and-contract.PRS-IND chest.PL-NOM-M.3PL.POSS in.the.face.of every-GEN

lit. it is worth mentioning that the scandal opened fire at the American vice-president

- (46) a. bala‘a l-ṭu‘m-a
 swallow.PST DEF-bait-ACC
 he swallowed the bait
- b. alqāhirat-u wa-dimashq-u lam tabtali‘-a
 Cairo-NOM and-Damascus-NOM NEG.PST swallow.F-DU.JUSS
 l-ṭu‘m-a (Al-Ahram, 1999)
 DEF-bait-ACC
 Cairo and Damascus did not swallow the bait

However, in the following two idioms, the verbs inflect fully for gender, person, and number, but they only occur in the present tense form in *Arabic corpus* and the *LDC Arabic corpus*:

- (47) a. yaḥruth-u fi l-baḥr-i
 plow.PRS-IND in DEF-sea-GEN
 he is plowing the sand
- b. jināyat-u-hu ‘anna-hu kāna yaḥruth-u
 crime-NOM-M.3SG.POSS COMP-M.3SG be.PST.M.3SG plow.PRS.M.3SG-IND
 fi l-baḥr-i (Al-Ahram, 1999)
 in DEF-sea-GEN
 lit. his crime is that he was plowing in the sea
- (48) a. yanfukh-u fī qirbat-i-n maqtū‘at-i-n
 blow.PRS-IND in bagpipe-GEN-INDF cut.PPTCP-GEN-INDF
 he is trying in vain to convince other people to take certain actions to improve a situation
- b. naḥnu na‘lam-u ‘anna-nā nanfukh-u fī
 1PL know.PRS.1PL-IND COMP-1PL blow.PRS.1PL-IND in
 qirbat-i-n maqtū‘at-i-n (Al-Thawra, after 2000)
 bagpipe-GEN-INDF cut.PPTCP-GEN-INDF
 lit. we know that we are blowing into a cut-off bagpipe

The idiomatic meanings can, however, be situated in the past by using the verbs in their present tense forms after the past tense marker *kāna* “was” or one of its inflectional forms, as in (47b). This structure communicates “the idea of continued or habitual action in the past” (Ryding 2005: 448).

The Arabic past tense form indicates that the action is completed. The actions expressed in the literal readings of both idioms cannot be completed, that is, the sea cannot be plowed and a bagpipe with a hole in it cannot be inflated to produce music. Indeed, these actions, if commenced, could *continue* forever without achieving their intended results. Therefore, it could be proposed that using these verbs in the past tense form would be contradictory to the idiomatic meanings, since such use could suggest that the impossible goals were achieved. Thus, this

restriction on inflection could be explained by both the grammatical meaning of the past tense form in Arabic and the way in which the figurative images are employed to convey the conventionalized idiomatic meanings.

In other words, using these verbs in the past tense form would violate one of the cognitive constraints on idiom variation, i.e. *compatibility*. This “is the most important constraint for the creation of systematic variants [and it] points to the requirement that any formal modifications of the literal scene must be in accordance with the idiomatic meaning” (Langlotz 2006: 221).

Another relevant phenomenon in this context is that the verbs in (49a) and (50a) typically occur in the present tense form. This is what their examples in *Arabic corpus* and the *LDC Arabic corpus* suggest:

- (49) a. *lam* *yuṣaddiq* *‘ayn-ay-hi*
 NEG.PST believe.M.3SG.JUSS eye-DU.ACC-M.3SG.POSS
 he did not believe his eyes
- b. *lam* *yuṣaddiq* *‘ayn-ay-hi* *‘indamā*
 NEG.PST believe.M.3SG.JUSS eye-DU.ACC-M.3SG.POSS when
shāhada *dhālika* *l-makhlūq-a* (Al-Thawra, after 2000)
 see.PST.M.3SG DEM DEF-creature-ACC
 he did not believe his eyes when he saw that creature
- (50) a. *lā* *taqūm-u* *li-shakhṣ/shay’* *qā’imat-u-n*
 NEG stand.up-IND for-sb/sth pillar-NOM-INDF
sb/sth is completely ruined (e.g. socially or professionally)
- b. *lam* *taqum* *li-l’ansār-i* *ba‘da-hā*
 NEG.PST stand.up.JUSS for-Al’ansār-GEN after-F.3SG
qā’imat-u-n (from the non-fiction section in *Arabic corpus*)
 pillar-NOM-INDF
 after that, the [Medinan] followers [of Prophet Mohammed] were ruined

As their examples in the corpora exhibit, these idioms normally occur in negative polarity contexts. When the context requires their meanings to be situated in the past, as in the examples provided above, the verbs may be used in the present tense form preceded by *lam*. This is a negative which can only be used with the present tense form of the verb to communicate negation and convey a meaning of past tense (see e.g. Ryding 2005 and Badawi *et al.* 2004).

The other possible way to convey past-time negation in Arabic is by using the negative *mā* followed by the verb in its past tense form. As Holes (2004: 323) points out:

[t]he original difference between these forms in CLA [...] lay not so much in their temporal or aspectual reference as in the degree of personal participation and “affect” in the action or state expressed by the negativized verb: the *ma*: [i.e. *mā*] construction was typically used for direct (especially first person) speech and the *lam* construction was more common in the description of third person action.

Elsewhere, Holes also describes the difference between the two constructions in CLA and comments on its status in MSA. He notes that the *mā* construction was used to negate completed actions:

This was particularly likely in first person direct speech where the speaker was certain of the truth of what he was saying [. . .], whereas there was a tendency for reports of third-party action, where the speaker might be less certain or less committed to the truth of the proposition, to be expressed by [the] *lam* [. . . construction]. Neither of these semantic-affective distinctions in the use of negative particles is any longer observed to the same extent or in the same way in contemporary MSA.

(Holes 2004: 239–240)

Ryding (2005) describes the use of *mā* construction as “rare in written Arabic, although it is widely used in spoken Arabic vernaculars” (p. 647). Holes (2004: 323–324) notes that this structure is widely employed in the spoken dialects and also occurs in literary writings:

[I]n the nonliterary written usage of journalism and other expository text, the *lam* construction has come to predominate overwhelmingly in all contexts where a past-time negative is required. The use of *ma:* construction has come to be associated either with a consciously “literary” style or with dialectal usage; on both accounts it is avoided in nonliterary MSA writing.

To him, this avoidance is an example of a general tendency to consciously distance “written Arabic from anything that smacks of dialect” (p. 324).

I have searched the *Modern Literature* and *Non-fiction* sections in *Arabicorpus* (see section 2.4.1.1) but have not found any examples of these two idioms used in the *mā* construction. However, I have found examples of them used with *mā* in some other texts that show features of spoken Arabic. The avoidance of this construction in (general) written Arabic may therefore explain why these negative polarity idioms are usually used in the present tense form in the data.

It should be noted, however, that the verb in (50a) has been used in the past tense form in two cases: first, when there is a grammatical necessity for the use of *mā*, e.g. when the idiom is used as the apodosis of a conditional sentence, as in example (102) below. Second, this verb has been used in the past tense form in some contexts where the meaning of negation is not expressed by an overt negative. For instance, depending on context, some rhetorical questions can express denial or give rise to, at least, very strong doubts about the possibility of the action conveyed by the verb. Tovena (1996) notes that “[t]he notion of negative contexts covers more than environments containing overt negation or negative quantifiers. Elements that induce a negative context are potential licensors for negative polarity sensitive items” (p. ii).

As far as the inflection of verbal idioms for person is concerned, some of these, e.g. (51) to (53), are more likely to occur in the first person than others, e.g. (54) to (56), are:

- (51) a. *ḍaraba* *‘uṣfūr-ayni* *bi-ḥajar-i-n*
hit.PST sparrow-DU.ACC.INDF with-stone-GEN-INDF
he killed two birds with one stone
- b. *wa-bi-hādhā* *naḍrib-u* *‘uṣfūr-ayni*
and-by-DEM hit.PRS.1PL-IND sparrow-DU.ACC.INDF
bi-ḥajar-i-n *wāḥid-i-n* (Al-Watan, 2002)
with-stone-GEN-INDF one-GEN-INDF
lit. and by this we hit two sparrows with one stone
- (52) a. *rafa‘a* *rāyat-a* *shay’*
raise.PST flag-ACC *sth*
he supported *sth*
- b. *‘arfa‘-u* *rāyat-a* *l-taghyīr-i*
raise.PRS.1SG-IND flag-ACC DEF-change-GEN
l-jizriyy-i (Al-Ahram, 1999)
DEF-radical-GEN
lit. I raise the flag of radical change
- (53) a. *ghariqa* *ḥattā* *‘uthun-ay-hi* *fi* *shay’*
sink.PST up.to ear-DU.GEN-M.3SG.POSS in *sth*
lit. he sank up to his ears in *sth*
he became deeply involved in *sth*
- b. *ghariq-tu* *fi* *qaḍiyyat-i* *l-sharq-i* *l-‘awsaṭ-i*
sink.PST-1SG in problem-GEN DEF-East-GEN DEF-middle-GEN
ḥattā *‘uthun-ay-y* (Al-Ahram, 1999)
up.to ear-DU.GEN-1SG.POSS
lit. I sank in the problem of the Middle East up to my [two] ears
- (54) *dassa* *l-summ-a* *fi* *l-‘asal-i*
put.secretly.PST DEF-poison-ACC in DEF-honey-GEN
he deceptively promoted something bad (e.g. a way of thinking) by presenting it in an advantageous or attractive way
- (55) *‘arghā* *wa-‘azbada*
foam.PST and-froth.PST
he fumed with rage
- (56) *iṣṭāda* *fi* *l-mā’-i* *l-‘akir-i*
fish.PST in DEF-water-GEN DEF-cloudy-GEN
he fished in troubled waters

The fact that examples (54) to (56), and the like in the data, typically express negative evaluations of their subjects may account for why they do not usually occur in the first person. Using them in this way would suggest self-abasement or attribute a socially unacceptable quality to the speaker or writer (see Strässler 1982). For this reason, the use of these examples in the first person would be expected only in contexts where the speaker or writer is not concerned about self-abasement or loss of face.

Even when such idioms do occur in the first person, they are usually negated, as in (57b), or deprived of their negative evaluative content through appropriate management of the context, as in (58b) where the typically negatively evaluative

idiom is used approvingly. However, in this particular example, the context does not completely rule out the literal interpretation of this metaphoric idiom:

- (57) a. 'adāra 1-khadd-a 1-'ākhar-a
 turn.PST DEF-cheek-ACC DEF-other-ACC
 he turned the other cheek
- b. sawfa lā nudīr-u khadd-a-na 1-'ayman-a
 FUT NEG turn.PRS.1PL-IND cheek-ACC-1PL.POSS DEF-right-ACC
 li-man yaḍrib-u-nā 'alā khadd-i-na
 to-REL hit.PRS.M.3SG-IND-1PL on cheek-GEN-1PL.POSS
 1-'aysar-i (Al-Hayat, 1996)
 DEF-left-GEN
 lit. we will not turn our right cheek to that who hits us on our left cheek
- (58) a. saraḥa wa-mariḥa
 move.freely.PST and-enjoy.oneself.PST
 he did as he liked
- b. kāna 'ālam-u 1-ṭufūlat-i bi-l-nisbat-i
 be.PST world-NOM DEF-childhood-GEN with-DEF-relation-GEN
 'ilayy firdaws-a-n 'ishtu khilāla-hu mutaḥarrir-a-n
 to.1SG paradise-ACC-INDF live.PST.1SG during-M.3SG free-ACC-INDF
 min-a 1-humūm-i 'asraḥ-u
 of DEF-worry.PL-GEN move.freely.PRS.1SG-IND
 wa-'amraḥ-u kamā shā'a li
 and-enjoy.oneself.PRS.1SG-IND as want.PST for.1SG
 llāh-u (Al-Hayat, 1996)
 Allah-NOM
 lit. to me, the world of childhood was a paradise, during which I lived free of worries, moving freely and enjoying myself, as Allah intended for me

As far as the inflection for the second person is concerned, idioms that convey negative evaluations of their subjects may only be “used of someone considered less powerful and therefore open to loss of face” (Moon 1998: 261). The verbal idioms in my data do not typically occur in the second person. An additional explanation for this is that *Arabicorpus* consists mainly of written texts. The use of the second person form is more typical of spoken language, as in this medium the interlocutors usually share the same space and time.

The foregoing analysis shows that some of the morphosyntactic rules regarding the use of idioms are not “hard” ones, i.e. they do not sanction one route and rule out any other possible routes. Rather, they are “soft” rules that can be expressed quantitatively (Fellbaum 2007). It shows as well that it is essential sometimes for formulating these rules to consider the properties of the situational context in which idioms are put to use, the medium of interaction, and/or the text type.

5.4.1.2 *Inflectability of nouns and adjectives*

The restrictions on inflectability are most noticeable in the nominal and adjectival components of idioms. In many cases, the nominal and adjectival elements in the data show no inflectability in their occurrences in *Arabic corpus*. Examples of these include the nouns and adjectives in:

- (59) qalaba zahr-a l-mijann-i li-*shakhš/shay*'
 turn.around.PST back-ACC DEF-shield-GEN to-*sb/sth*
 he became hostile to *sb/sth* after he was a friend or supporter of them/it
- (60) *shakhš/shay*'_i bi-rummat-i-*hi*_i
*sb/sth*_i with-worn.out.piece.of.rope-GEN-M.3SG.POSS_i
 all *sb*_i (e.g. a group of people) or the whole *sth*_i
- (61) *shakhš/shay*'_i 'alā bakr-at-i 'ab-ī-*hi*_i
shakhš/shay'_i on young.camel-F-GEN father-GEN-M.3SG.POSS_i
 all *sb/sth*_i (e.g. a group of people or things)
- (62) nār-u-n 'alā 'alam-i-n
 fire-NOM-INDF on mountain-GEN-INDF
 very well-known entity
- (63) bala'a l-ṭu'm-a
 swallow.PST DEF-bait-ACC
 he swallowed the bait
- (64) 'amsaka l-'aṣā min-a l-muntaṣaf-i
 hold.PST DEF-stick at DEF-middle-GEN
 he struck a balance
- (65) yanfukh-u fī qirbat-i-n maqtū'at-i-n
 blow-IND into bagpipe-GEN-INDF cut.PPTCP-GEN-INDF
 he is trying in vain to convince other people to take certain actions to improve a situation
- (66) 'ibrat-u-n fī kawmat-i qashsh-i-n
 needle-NOM-INDF in stack-GEN hay-GEN-INDF
 a needle in a haystack
- (67) qimmat-u jabal-i l-jalīd-i
 top-NOM mountain-GEN DEF-ice-GEN
 the tip of the iceberg
- (68) *lā* taqūm-u li-*shakhš/shay*' qā'imat-u-n
 NEG stand.up.PRS.F-IND for-*sb/sth* pillar.F-NOM-INDF
sb/sth is socially or professionally ruined

Several factors that act individually or in different combinations could explain these restrictions. These are the nature of the figurative image on which the idiom is based, the semantics of the idiom, the transparency of the motivation of the idiom, and the logic of the grammatical environments in which the idiom occurs.

For example, the properties of the figurative image may explain the restrictions on the inflectability of the nominal component in the idiom in (65). That is, one

would normally blow into one bagpipe (or skin) at a time, this might account for why *qirbah* only occurs in the singular form.

Both the idiomatic meaning and the properties of the figurative image may explain why *qimmat* “top” in (67) only occurs in the singular form. The idiom refers to a small part of a much larger problem or phenomenon. The singularity of the head noun thus may be more suitable for the meaning “small” in the idiomatic signification. In addition, a mountain or an iceberg has (or is perceived to have) only one summit or tip, which could constitute a further account for the restriction.

The opaqueness and idiosyncrasy of the idioms in (59), (60) and (61) are obvious explanations for the restrictions on the inflectability of their nominal components.

Finally, in (68), the interaction of the figurative image and the logic of the grammatical environment in which the idiom occurs may account for the restriction on the inflectability of the noun, i.e. the fact that *qā'imah* “pillar” always occurs in the singular form. An image of a building (or perhaps a tent) that has none of its pillars up seems to be the one on which the figuration in the idiom is based. In order to convey its meaning, the idiom must occur within negative polarity contexts, as noted earlier. Within the scope of a negative, the singular indefinite noun is interpreted as expressing the lack of even the least amount or lowest degree of its referent. This may account for why *qā'imah* does not occur in the dual or plural forms. It also explains the illogicality,[†] or at least the dubious status, of the following invented example:

- (69) *lam taqum li-l-farīq-i qawā'im-u ba'da dhālika
 NEG.PST stand.up.JUSS for-DEF-team-GEN pillar.PL-NOM after DEM
 lit. the pillars of the team did not stand up after that

in the meaning “none of the team’s pillars stood up after that”.

However, the nominal and adjectival elements in many other idioms show some inflectability in the corpus texts, e.g.:

- (70) ba'īd-u l-ghawr-i
 far-NOM DEF-bottom-GEN
 e.g. (of emotions) deeply felt or (of changes) profound
- (71) tajammada l-dam-u fī 'urūq-i shakhṣ
 freeze.PST DEF-blood-NOM in vein.PL-GEN sb
 the blood froze in sb’s veins
- (72) qaraba 'uṣfūr-ayni bi-ḥajar-i-n
 hit.PST sparrow-DU.ACC.INDF with-stone-GEN-INDF
 he killed two birds with one stone
- (73) shakhṣ_i/shay' ṭawwaqa 'unuq-a shakhṣ_j
 sb_i/sth surround.PST neck-ACC sb_j
 lit. sb_i/sth surrounded sb_j's neck
 sb_i or sth (e.g. trust) greatly honored sb_j, or sth (e.g. debt) burdened sb_j
- (74) dam-u-n 'azraq-u
 blood-NOM-INDF blue-NOM
 blue blood

In addition to this type, Badawi *et al.* (2004) point out that the verbs *tamma* “to complete”, *jarā* “to proceed”, and *qāma bi-* “to carry out” are frequently used to paraphrase a passive (pp. 431ff). In this case, the verbal nouns of the verbs that need to be passivized are used as the heads of the subject phrases of these verbs. The verb *tamma*, for example, may be translated in this structure into *be done*, *completed*, or *carried out*, for instance, depending on context.

In his account of the phenomenon, Holes (2004: 317) points out that:

[t]his [periphrastic] structure is used instead of an internally voweled passive [i.e. a formal passive] to report the completion of durative or iterative processes (rather than for the description of punctual events, cognitive activities, and emotional states) where the focus is on the result of the process rather than on the process itself or on the (usually multiple rather than individual) agents who performed it.

However, it should be noted that both types of passive, as impersonal structures, normally shift the focus from the agents involved. For example, Ryding (2005) states that the passive is “used in Arabic only if the agent or doer of the action is non-designated, unknown, or not to be mentioned for some reason” (p. 658). Badawi *et al.* (2004) also mention that “Arabic passive is intrinsically impersonal and hence often used for impersonal and vague constructions in English” (p. 388).

Afonso (2008: 105) lists the following reasons for the suppression of the agent in impersonal constructions in general:

The agent may be demoted or suppressed because the speaker i) is unaware of the identity of the agent, ii) is aware but does not wish to reveal his/her identity, iii) is aware of the identity and assumes the hearer is also aware, making the reference to the agent obvious and its expression redundant, and iv) is aware of the identity of the agent but considers it unimportant.

In some cases, the patient of the action or the action itself may be viewed by the speaker or writer as more important than the agent. Thus, using the passive would help set the attention of the hearer or reader on the most important piece of information from the point of view of the speaker or writer.

Having said that, MSA has been undergoing some changes with this respect. For example, Badawi *et al.* (2004) note that the agent in the case of the passive construction “cannot in theory be mentioned even periphrastically elsewhere in the sentence, though MWA [i.e. Modern Written Arabic] is starting to do so under the influence of European languages” (p. 383).

The discussion of the active and passive voice in the context of this study is only relevant to verbal idioms. The following questions are addressed: How can the base forms of these idioms be classified regarding their voice? Do verbal idioms allow for any changes in their voice? In what ways is the passive voice expressed in Arabic idioms? And, what function(s) does the passive voice have when it occurs?

The base forms of all verbal idioms in the data are in the active voice form. The formal passive occurs in four verbal idioms. They are:

- (76) a. rafa‘a rāyat-a *shay’*
 raise.PST flag-ACC *sth*
 he supported *sth*
- b. ‘inna-hu la-shay’-u-n gharīb-u-n ḥaqqan
 indeed-M.3SG EMPH-thing-NOM-INDF strange-NOM-INDF really
 ‘an turfa‘-a rāyāt-u l-riddat-i
 COMP raise.PRS.F.PASS-SBJV flag.F.PL-NOM DEF-retrogression-GEN
 l-‘ilmiyyat-i fī ‘aṣr-i-n tatasābaq-u fī-hi
 DEF-scientific-GEN in age-GEN-INDF race.PRS-IND in-M.3SG
 l-‘umam-u naḥwa l-‘akhdh-i
 DEF-nation.PL-NOM towards DEF-taking-GEN
 bi-‘ulūm-i l-ḥāsib-i (Al-Ahram, 1999)
 with-science.PL-GEN DEF-computer-GEN
 lit. indeed, it is really a strange thing that the flags of scientific retrogression are raised in an age in which all nations are racing to acquire computer sciences
- (77) a. waḍa‘a *shakḥ/shay’* ‘ala l-raff-i
 put.PST *sb/sth* on DEF-shelf-GEN
 he made *sb/sth* no longer useful or desirable
- b. ‘inna khuttāt-a l-ta‘wīm-i fī l-sūq-i
 indeed plan-ACC DEF-floating-GEN in DEF-market-GEN
 wuḍi‘a-t ‘ala l-raff-i (Al-Hayat, 1996)
 put.PST.PASS-F on DEF-shelf-GEN
 lit. indeed, the plan of floating [the exchange rate] in the market was put on the shelf
- (78) a. ṣabba l-mā’-a l-bārid-a ‘alā *shay’*
 pour.PST DEF-water-ACC DEF-cold-ACC on *sth*
 he poured cold water on *sth*
- b. fa-kullamā kāna hunāka tawajjuh-u-n nashīt-u-n
 and.so-every.time be.PST there aiming-NOM-INDF active-NOM-INDF
 li-taḥrīk-i mawqif-i-n ‘aw sa‘y-u-n
 at-getting.sth.moving-GEN situation-GEN-INDF or effort-NOM-INDF
 naḥwa taf‘īl-i ‘amal-i-n mushtarak-i-n lā
 towards activating-GEN work-GEN-INDF joint-GEN-INDF NEG
 nalbath-u ‘illā qalīl-a-n thumma yuṣabb-u
 stay.PRS.1PL-IND but little-ACC-INDF then pour.PRS.PASS-IND
 ‘alayhi mā’-u-n bārid-u-n (Al-Ahram, 1999)
 on.M.3SG water-NOM-INDF cold-NOM-INDF
 lit. and so every time there is active aiming at getting a situation moving or an effort to activate joint work, it does not take long before cold water is poured on it

- (79) a. naza‘a fatīl-a *shay’*
 pull.out.PST fuse-ACC *sth*
 he defused *sth* (e.g. a dangerous situation)
- b. wa-man kāna yataṣawwar-u ’anna l-ṣirā‘-a l-‘arabiyy-a
 and-Q be.PST imagine-IND COMP DEF-conflict-ACC DEF-Arab-ACC
 l-‘isrā’īliyy-a yunza‘-u min-hu fatīl-u
 DEF-Israeli-ACC pull.out.PRS.PASS-IND from-M.3SG fuse-NOM
 l-ṣidām-i l-musallaḥ-i li-yataḥawwal-a ’ilā
 DEF-clash-GEN DEF-armed-GEN to-transform-SBJV to
 tafāwuḍ-i-n (Al-Ahram, 1999)
 negotiation-GEN-INDF
 lit. who was able to imagine that the Arab-Israeli conflict (subject of
 clause) the fuse of armed clash is pulled out from it to become negotia-
 tion (predicate)

In four idioms, the passive meaning has been conveyed through the use of *tamma* “to complete” plus the verbal noun form of the idiomatic verb, e.g.:

- (80) a. iṣṭāda fi l-mā’-i l-‘akir-i
 fish.PST in DEF-water-GEN DEF-cloudy-GEN
 he fished in troubled waters
- b. ḥattā lā yatimm-a l-iṣṭiyād-u fi l-mā’-i
 in.order.that NEG be.done.PRS-SBJV DEF-fishing-NOM in DEF-water-GEN
 l-‘akir-i (Al-Watan, 2002)
 DEF-cloudy-GEN
 lit. in order that fishing in cloudy water is not done
 so that no one would fish in troubled waters
- (81) a. naza‘a fatīl-a *shay’*
 pull.out.PST fuse-ACC *sth*
 he defused *sth* (e.g. a dangerous situation)
- b. wa-bi-hādhīhi l-istiḳālat-i sa-tatafajjar-u
 and-by-DEM DEF-resignation-GEN FUT-explode-IND
 ’azamāt-u-n kathīrat-u-n rubbamā yatimm-u
 crisis.PL.F-NOM-INDF many.F-NOM-INDF perhaps complete.PRS-IND
 naz‘-u faṭīl-i-hā bi-’i‘lān-i
 pulling.out-NOM fuse-GEN-F.3SG.POSS by-announcing-GEN
 l-tashkīl-i l-ḥukūmiyy-i l-jadīd-i
 DEF-formation-GEN DEF-governmental-GEN DEF-new-GEN
 (Al-Ahram, 1999)
 lit. and by this resignation, many crises [that] the pulling out of their fuse
 may be done by announcing the new formation of the cabinet (subject of
 clause) will explode (predicate)

All the examples above are isomorphic idioms. In the cases where they feature objects among their components, their isomorphism could explain why the

objects can be moved to the position of the topic of the sentence in the passivized versions.

Given the account provided above for the preference of the periphrastic passive over the formal one or vice versa, in many cases, selecting one of them rather than the other seems to be a matter of choice that depends on contextual factors rather than on formal ones. This may explain why the same idiom, e.g. in (79) and (81), can be used in both ways.

Having said that, this point, among some other points in this chapter, would be best examined on the basis of more sets of data from different sources, perhaps with more use of appropriate interviews with native informants, in order to determine if and in what way(s) they would passivize these idioms.

The corpus examples in (78b) and (80b), particularly, negatively evaluate the unstated agents of the idioms. The use of the passive here then helps the writer hide the identity of the agent in order not to be direct in expressing their negative judgment. As the context of (78b) shows, the writer disapproves of some (Arab) politicians who, they claim, stand in the way of any effective cooperation among the Arab countries.

In (80b), the context shows that the writer approves of the supervision of the government on charity work in Kuwait. He also anonymously refers to, and disapproves of, some people who want this work to become under the control of different social and religious groups and tribes in the country. According to him, they only call for this change because it would help them “fish in troubled waters”, i.e. benefit from the possibly bad relationships between these groups, and gain advantages in their election campaigns.

It is worth mentioning here too that in many cases in the corpus, the agent of this idiom is not explicitly stated in the text even when the verb is used in the active voice form. For example, the agent is sometimes referred to by words such as *al-ba‘d* “some [people]” and *alladhīna* “those who”. This again helps the writer be somewhat vague by avoiding directing the negative evaluation to a specified entity. However, the question of what types of lexical item collocate with and, particularly, fill the structural slots in the idioms may be an interesting area for future research on the subject. This question has only been touched upon in the present study when need arises.

Avoiding direct assignment of negative judgments to the agent is useful, for instance, when the agent is a powerful entity that the speaker or writer does not want to be involved in a confrontation with or when the speaker or writer enjoys good relationships with the agent which they do not want to risk. By using the passive, the responsibility for determining the agent lies with the hearer or reader, who is forced to rely heavily on their knowledge of the (prior) context so as to decode the message. This may be a good strategy for the speaker or writer to hide behind and deny responsibility for the hearer’s or reader’s interpretation of their own text.

The relatively low incidence of this transformation in the data might be explained by the fact that the passivized constructions are usually highly marked in text.

As opposed to all the verbal idioms in the data, the following idiom, which is not in the random sample, occurs in the passive voice in its base form:

and Hunston 2001). Even more, in some cases, people may direct their, particularly, negative evaluations towards each other's behavior as an avoidance strategy, that is, in order to conceal their wish to negatively evaluate each other, because such an evaluation may appear less objective or, even more importantly, could be regarded as motivated by malice or hatred. When such actions are denoted, and frequently evaluated, through idioms, they are typically expressed using verbal groups.

Tables 5.1 to 5.6 show the syntactic structures of Arabic idioms. This information has been given mainly in reference to the base forms of the idioms. In each table, the first column contains formulae that show the syntactic structures of the idioms and the grammatical functions of their higher-level constituents. As long as their status is obvious, the grammatical functions of the lower-level constituents have not been represented in the tables. The second column gives information on the frequency of the specified structure in the data. Finally, the third column provides a relevant example from the data. The structures are ordered according to their frequency within each category.

In the first column, items given in italics do not represent any idiomatic components. Rather, they typically show the positions of the syntactic slots that need to be filled in by elements from the context. However, information on the position of the subject slot in the case of verbal idioms has been given only when this has been considered essential for clarity, e.g. when the subject is co-referential with a possessive that is supplied by the context.

Table 5.1 The syntactic structures of verb-subject idioms

<i>Syntactic structure and grammatical functions</i>	<i>F.</i>	<i>Example</i>
V [_{SBJ} NP[N NP]]	3	'afala najm-u <i>shakhṣ/shay'</i> set.PST star-NOM <i>sb/sth</i> the glory or fame of <i>sb/sth</i> ended
<i>NP</i> _i V [_{SBJ} NP[N <i>POSS</i> _i]] [_{COMP} PP[P NP[N N]]]	1	<i>shay'</i> 'akhadha ba'd-u- <i>hu</i> <i>sth</i> hold.on.PST some-NOM-M.3SG.POSS bi-riqāb-i ba'd to-neck.PL-GEN some <i>sth</i> (e.g. a set of ideas, texts, or problems) interconnected
V [_{ADJUNCT} PP[P NP]] [_{SBJ} N]	1	<i>lā</i> taqūm-u li- <i>shakhṣ/shay'</i> NEG stand.up.PRS.F-IND for- <i>sb/sth</i> qā'imat-u-n pillar.F-NOM-INDF <i>sb/sth</i> is completely ruined e.g. socially or professionally
V [_{SBJ} NP[DET N]] [_{ADJUNCT} PP[P NP[N NP]]]	1	tajammada l-dam-u fī freeze.PST DEF-blood-NOM in 'urūq-i <i>shakhṣ</i> vein.PL-GEN <i>sb</i> the blood froze in <i>sb</i> 's veins

Table 5.2 The syntactic structures of verbal idioms

Syntactic structure and grammatical functions	F.	Example
V [OBJ NP[N NP]]	7	rafa‘a rāyat-a shay’ raise.PST flag-ACC sth he supported sth
V [OBJ NP[DET N]]	3	bala‘a l-tu‘m-a swallow.PST DEF-bait-ACC he swallowed the bait
V [ADJUNCT PP[P NP[N NP]]]	2	taqallaba ‘alā jamr-i shay’ turn.over.PST on ember.PL-GEN sth he severely suffered from sth
V [OBJ NP[DET N]] [ADJUNCT PP[P NP [DET N]]]	2	‘amsaka l-‘aṣā min-a l-muntaṣaf-i hold.PST DEF-stick at DEF-middle-GEN he struck a balance
V CONJ V	2	‘arghā wa-‘azbada foam.PST and-froth.PST he fumed with rage
V [ADJUNCT PP[P NP[DET N]]]	1	yaḥruth-u fi l-baḥr-i plow.PRS-IND in DEF-sea-GEN he is plowing the sand
V [OBJ NP[N NP[DET N]]] [ADJUNCT PP[P NP]]	1	qalaba zahr-a turn.around.PST back-ACC l-mijann-i li-shakḥ/shay’ DEF-shield-GEN to-sb/sth he became hostile to sb/sth after he was a friend or supporter of them/it
V [OBJ NP[N NP[DET N]]]	1	ḥāraba ṭawāḥīn-a l-hawā’-i fight.PST mill.PL-ACC DEF-air-GEN he tilted at windmills
V [OBJ NP[NP[DET N]] AP[DET ADJ]]]	1	‘adāra l-khadd-a l-‘ākhar-a turn.PST DEF-cheek-ACC DEF-other-ACC he turned the other cheek
NP _i V [OBJ NP[N NP [N POSS _i]]]	1	‘akmala niṣf-a dīn-i-hi perfect.PST half-ACC religion-GEN-M.3SG.POSS he got married
V [ADJUNCT PP[P NP[NP [DET N] AP[DET ADJ]]]]	1	iṣṭāda fi l-mā’-i l-‘akir-i fish.PST in DEF-water-GEN def-cloudy-GEN he fished in troubled waters
NP _i V [OBJ NP[N NP _j]] [ADJUNCT PP[P NP[N POSS _j]]]	1	ḥafara qabr-a shakḥ/shay’ dig.PST grave-ACC sb/sth bi-yad-i-hi with-hand-GEN-M.3SG.POSS he put an end to e.g. the glory or success of sb, or he put an end to sth

(Continued Overleaf)

Table 5.2 Continued

<i>Syntactic structure and grammatical functions</i>	<i>F.</i>	<i>Example</i>
V [OBJ NP[DET N]] [ADJUNCT PP[P NP]]	1	fataḥa l-nār-a ‘alā <i>shakḥ/shay</i> open.PST DEF-fire-ACC at <i>sb/sth</i> he verbally attacked <i>sb/sth</i>
V [OBJ NP] [ADJUNCT PP[P NP[DET N]]]	1	waḍa‘a <i>shakḥ/shay</i> ‘ala l-raff-i put.PST <i>sb/sth</i> on DEF-shelf-GEN he made <i>sb/sth</i> no longer useful or desirable
V [ADJUNCT PP[P NP[N ADJ]]]	1	yanfukh-u fī qirbat-i-n blow-IND into bagpipe-GEN-INDF maqtū‘at-i-n cut.PPTCP-GEN-INDF he is trying in vain to convince other people to take certain actions to improve a situation
V [ADJUNCT PP[P NP[N POSS]]] [ADJUNCT PP[P NP]]	1	ghariqa ḥattā ‘uthun-ay-hi sink.PST up.to ear-DU.GEN-M.3SG.POSS fī <i>shay</i> in <i>sth</i> he was up to his ears in <i>sth</i>
REL _i V [ADJUNCT NP[N POSS _i]]	1	<i>shakḥ/shay</i> _i wa-man/mā <i>sb/sth</i> _i and-REL.HUM/REL.NONHUM laffa laff-a- <i>hu</i> _i turn.round.PST turning.round-ACC-M.3SG.POSS _i <i>sb</i> _i and that who did just as they did, <i>sb</i> _i and that who is of their type, or <i>sth</i> _i and that which is of its kind
V [ADJUNCT NP[N NP[DET N]]] [ADJUNCT PP[P NP]]	1	<i>shay</i> _i marra murūr-a <i>sth</i> pass.PST passing-ACC l-kirām-i bi- <i>shakḥ</i> DEF-noble.people-GEN by- <i>sb</i> <i>sth</i> passed unnoticed by <i>sb</i>
V [OBJ N] [ADJUNCT PP[P N]]	1	ḍaraba ‘uṣfūr-ayni hit.PST sparrow-DU.ACC-INDF bi-ḥajar-i-n with-stone-GEN-INDF he killed two birds with one stone
V [OBJ NP[NP[DET N] AP[DET ADJ]]] [ADJUNCT PP[P NP]]	1	ṣabba l-mā‘-a l-bārid-a pour.PST DEF-water-ACC DEF-cold-ACC ‘alā <i>shay</i> on <i>sth</i> he poured cold water on <i>sth</i>

Table 5.3 The syntactic structures of nominal idioms

<i>Syntactic structure and grammatical functions</i>	<i>F.</i>	<i>Example</i>
[_{HEAD} N] [_{MOD} ADJ]	5	rimāl-u-n mutaḥarrikat-u-n sand.PL-NOM-INDF move.APTCP-NOM-INDF quicksand
[_{HEAD} N] [_{MOD} NP[DET N]]	3	ghuṣn-u l-zaytūn-i branch-NOM DEF-olive-GEN olive branch
[_{HEAD} N] [_{MOD} N]	2	qarṣat-u ʾudhun-i-n pinch-NOM ear-GEN-INDF a light punishment that serves as a warning of a possible cruel one
[_{HEAD} N] [_{MOD} PP[P N]]	2	ḥibr-u-n ʿalā waraq-i-n ink-NOM-INDF on paper-GEN-INDF inactive (e.g. of decisions or laws)
[_{HEAD} N] [_{MOD} PP[P NP[N N]]]	1	ʾibrat-u-n fī kawmat-i qashsh-i-n needle-NOM-INDF in stack-GEN hay-GEN-INDF a needle in a haystack
[_{HEAD} N] [_{MOD} NP[N NP[DET N]]]	1	qimmat-u jabal-i l-jalīd-i top-NOM mountain-GEN DEF-ice-GEN the tip of the iceberg
[_{HEAD} NP[DET N]] [_{MOD} AP[DET ADJ]]	1	al-yad-u l-ʿulyā DEF-hand-NOM DEF-upper the upper hand
N CONJ N	1	shadd-u-n wa-jadhb-u-n tugging-NOM-INDF and-pulling-NOM-INDF great tension or struggle (e.g. in a strained political or social state)
REL CONJ REL	1	allutayyā wa-llaṭī REL.F.SG.DIMIN and-REL.F.SG much ado, or lengthy difficult discussions

Often, the tables show that the most frequent structures in a syntactic category tend to be its grammatically simplest ones. For instance, the idiomatic components of the most frequent structure in verb-subject idioms contain a verb and a subject. In verbal idioms, the most frequent structures contain a verb and an object with or without a determiner. In nominal idioms, the most frequent structures consist mainly of a head noun and a modifying element which realizes as an adjective or a noun.

One plausible, partial explanation may be suggested, in particular, for the relatively high frequency of the N + Adj structure in nominal idioms. That is, all the six idioms featuring this structure, given in (84) to (89) below, are isomorphic and convey both information and evaluation. The informational content in these cases consists in an entity or a situation that has a particular characteristic. Typically,

Table 5.4 The syntactic structures of prepositional idioms

<i>Syntactic structure and grammatical functions</i>	<i>F.</i>	<i>Example</i>
P NP[N NP]	4	fī daw'-i shay' in light-GEN sth in the light of sth
P NP[DET N]	2	taḥta l-'adwā'-i under DEF-light.PL-GEN in the spotlight
P NP[N N]	1	'alā marmā hajar-i-n at distance.that.sth.is.thrown stone-GEN-INDF of proximity in space or time
P NP[N ADJ]	1	'alā ṣafīḥ-i-n sākhin-i-n on sheet.iron-GEN-INDF hot-GEN-INDF (of a situation) very tense or instable, or (of a person) very anxious
P N	1	bayna nār-ayni between fire-DU.GEN-INDF between the devil and the deep blue sea
NP _i P NP[N N POSS _i]	1	shakhs/shay' _i 'alā bakr-at-i sb/sth _i on young.camel-F-GEN 'ab-ī-hi _i father-GEN-M.3SG.POSS _i all sb/sth _i (e.g. a group of people or things)

Table 5.5 The syntactic structures of adjectival idioms

<i>Syntactic structure and grammatical functions</i>	<i>F.</i>	<i>Example</i>
ADJ NP[DET N]	4	ba'īd-u l-ghawr-i far-NOM DEF-bottom-GEN e.g. (of a feeling) deeply felt or (of a change) profound

Table 5.6 The syntactic structures of adverbial idioms

<i>Syntactic structure and grammatical functions</i>	<i>F.</i>	<i>Example</i>
N PP[P N]	2	janb-a-n 'ilā janb-i-n side-ACC-INDF by side-GEN-INDF side by side

nouns convey entities and adjectives denote qualities. Therefore, this syntactic structure seems to be the most suitable one to convey this general type of meaning.

One may ask then: why is this type of meaning frequent in the data? Given the form-meaning connection referred to above, a good answer to this question could provide in turn a plausible account for the apparent need to rely more on the N + Adj structure in producing nominal idioms.

Once more, it can be proposed that a plausible answer lies in the evaluative function of these idioms. It is common in daily life that entities and states of affairs that have particular qualities are subject to evaluation, whether positive or negative. When these qualities are adequately intense and/or when the situations recur repeatedly, they may become candidates for idiomatic reference and often realize in the N + ADJ structure:

- | | | |
|------|--|---|
| (84) | baqarat-u-n
cow-NOM-INDF
milch cow | ḥalūb-u-n
milch-NOM-INDF |
| (85) | 'uṣṭuwānat-u-n
record-NOM-INDF
repetitious boring talk about a particular topic | mashrūkhat-u-n
crack.PPTCP-NOM-INDF |
| (86) | ḥalqat-u-n
ring.F-NOM-INDF
vicious circle | mufraqh-at-u-n
cast.in.a.mold.PPTCP-F-NOM-INDF |
| (87) | dam-u-n
blood-NOM-INDF
blue blood | 'azraq-u
blue-NOM |
| (88) | rimāl-u-n
sand.PL.F-NOM-INDF
quicksand, or an unreliable basis for <i>sth</i> (e.g. a plan or project) | mutaḥarrik-at-u-n
move.APTCP-F-NOM-INDF |
| (89) | al-yad-u
DEF-hand.F-NOM
the upper hand | l-'ulyā
DEF-upper.F |

5.4.3.2 Constituent order

Holes (2004) outlines the general grammatical structure of the sentence in MSA as follows:

[A] sentence in written Arabic consists of a subject and predicate. The subject (= S) may be freestanding, that is, a noun/independent pronoun; or dependent, that is, consisting of one or more bound morphemes that form part of the verb (= V) (if there is one) and that indicate the person, number and gender of the subject. The predicate may or may not contain a verb. If it does contain one, the subject may or may not be freestanding; if it does not, the sentence subject must be freestanding. The verb may or may not have a complement (= COMP) (p. 251).

There are some restrictions on the constituent order in Arabic. For example, MSA does not usually allow for the subject of the main clause to be fronted if this subject is indefinite (see e.g. Ryding 2005 and Badawi *et al.* 2004). But when there is no grammatical necessity for a particular order, different arrangements of the same elements may occur, depending on the involvement of some specific factors. However, Holes (2004: 251) points out that the verbal sentence in MSA has a normal order:

that is determined primarily by a principle of information organisation: what is already “known” from the previous text or context (and is usually grammatically definite) precedes what is “new” (and is usually indefinite), regardless of whether what is known/definite is the grammatical subject or object. In cases where both subject and object are known, or both are new, the subject, all things being equal, precedes the object. In MSA [. . .], these basic principles account for word order in the vast majority of sentences in prose texts, but may be overridden in cases of contrastive emphasis or topicalisation [. . .]

Foregrounding or preposing a constituent often lays emphasis on or gives more importance to its meaning vis-à-vis the meanings of the other parts in the same clause. For example, a speaker may choose to start with an element because they consider it to be more important than other elements, or when they are being challenged and think that there is a need to emphasize a specific element in the context in response to the challenge.

In addition to the principle of information organization and the need to place emphasis on or draw attention to some elements by preposing them in the sentence structure, the end-weight principle takes part in determining the constituent order in some cases. According to this principle, lexically “heavy” constituents generally follow “lighter” ones (see Holes 2004: 206f, for a discussion of this principle in MSA).

Fifteen idioms have shown variation in the constituent order of their base forms. These include three verb-subject idioms, ten verbal idioms, and two nominal idioms. The changes in these cases can be grouped under three headings: positioning the idiomatic subject, positioning the clause adjuncts, and coordinate reversal.

5.4.3.2.1 POSITIONING THE IDIOMATIC SUBJECT

Obviously, this point is only relevant to the six verb-subject idioms in the data. According to Badawi *et al.* (2004), MSA is a VSO language (p. 344). Holes (2004) also refers to the fact that this order is the most frequent one in the structurally simplest type of verbal sentence.

On the other hand, the SVO order is often used to draw attention to the subject (Ryding 2005). For this reason, among other possible ones, it is the order typically used in Arabic newspaper headlines. In actual fact, for a number of reasons, including the differences in the inflectional behavior of the verb in both orders, some accounts of constituent order in Arabic consider the SVO sentence, the *topic-comment sentence* in their terms, a distinct type of Arabic sentence that has strong

resemblance to *topicalization* in Western languages (see Badawi *et al.* 2004: 326). Also, on the functions of this order in MSA, Holes (2004: 252–253) notes that:

[e]xpository writing, which often involves defining key terms or describing the structure of entities with component parts, tends to contain a high proportion of sentences with SVCOMP order. [. . . In this case,] the text is “entity-oriented” [. . .] rather than “event-oriented”.

Two of the verb-subject idioms in the data have occurred in both the VS and SV orders:

- (90) a. 'afala najm-u *shakhs/shay'* or najm-u *shakhs/shay'* 'afala
 set.PST star-NOM *sb/sth* star-NOM *sb/sth* set.PST
 the fame or glory of *sb/sth* ended
- b. 'afala najm-u l-nāshir-i l-thariyy-i
 set.PST star-NOM DEF-publisher-GEN DEF-rich-GEN
 stīf fūrbis (Al-Hayat, 1996)
 Steve Forbes
 lit. the star of the rich publisher, Steve Forbes, set
- c. wa-l-najm-u 'afal (Al-Hayat, 1996)
 and-DEF-star-NOM set.PST
 lit. and the star set
- (91) a. dāra-t raḥā *shay'* or raḥā *shay'* dāra-t
 spin.PST-F quern *sth* quern *sth* spin.PST-F
sth (e.g. a war or elections) took place in a violent way
- b. wa-'akkada 'anna ma'ārik-a 'iddat-a-n
 and-stress.M.3SG.PST COMP DEF-battle.PL-ACC several-ACC-INDF
 dāriyat-a-n tadūr-u raḥā-hā bayna
 DEF-ferocious-ACC spin.F.PRS-IND quern-F.3PL.POSS between
 quwwāt-i-hi wa-bayna l-jaysh-i
 troop.PL-GEN-M.3SG.POSS and-between DEF-army-GEN
 l-sūdāniyy-i (Al-Hayat, 1996)
 DEF-Sudanese-GEN
 lit. and he stressed that several ferocious battles their quern is spinning
 between his troops and the Sudanese Army
- c. 'inna raḥa l-ma'rakat-i l-'askariyyat-i tadūr-u
 indeed quern DEF-battle-GEN DEF-military-GEN spin.PRS-IND
 fawqa l-'arḍ-i l-'afghāniyyat-i ḥaqqan
 on DEF-land-GEN DEF-Afghani-GEN really
 lākinna maqāšid-a-hā taqif-u ba'īdan (Al-Watan, 2002)
 but goal.PL-F.3SG.POSS stand.PRS-IND far.away
 lit. indeed, the quern of the military battle is really spinning on the
 Afghani land, but its goals stand far away

The other four verb-subject idioms have occurred only in the VS order:

- (92) inqabaḍa ṣadr-u *shakhṣ*^{||}
 contract.PST chest-NOM *sb*
sb felt depressed
- (93) tajammada l-dam-u fī ‘urūq-i *shakhṣ*
 freeze.PST.REFL DEF-blood-NOM in vein.PL-GEN *sb*
 the blood froze in *sb*'s veins
- (94) *lā* taqūm-u li-*shakhṣ/shay*' qā'imat-u-n
 NEG stand.up.PRS.F-IND for-*sb/sth* pillar.F-NOM-INDF
sb/sth is completely ruined e.g. socially or professionally
- (95) *shay*' 'akhadha ba'ḍ-u-*hu* bi-riqāb-i ba'ḍ
sth hold.on.PST some-NOM-M.3SG.POSS to-neck.PL-GEN some
sth (e.g. a set of ideas, texts, or problems) interconnected

The restriction on the SV order in (92) to (95) may be attributed to the emphasizing or attention-drawing functions of this order in MSA. Emphasizing or getting attention to a component of an expression normally entails that it bears an identifiable part of its meaning. The examples in (92) and (93) are metonymy-based idioms and the examples in (94) and (95) are metaphoric, and, therefore, all these four idioms are non-isomorphic, i.e. their components do not bear identifiable parts of their non-literal meanings. Thus, it is unlikely that a discursive need to get attention to or emphasize their subjects exists. This could explain why they have not occurred in the SV order. In addition, another syntax-related explanation is clearly relevant in the case of the idiom in (94). That is, the subject in this idiom cannot be fronted because it is indefinite.

On the other hand, the two idioms allowing the SV order, examples (90) and (91), are metaphorical and isomorphic, wherein their subjects bear parts of their non-literal significations and, therefore, can undergo syntactic operations that lay emphasis on or draw attention to them.

5.4.3.2.2 POSITIONING THE CLAUSE ADJUNCTS

Eighteen idioms in the sample contain adjunct phrases in their base forms. These include two verb-subject idioms and sixteen verbal idioms. As to the syntactic class of these adjuncts, two of them realize as noun phrases and the other sixteen as prepositional phrases. The nominal adjunct phrases occur in:

- (96) *shay*' marra murūr-a l-kirām-i bi-*shakhṣ*
sth pass.PST passing-ACC DEF-noble.people-GEN by-*sb*
sth (e.g. an event or anniversary) passed unnoticed by *sb*
- (97) *shakhṣ/shay*'_i wa-man/mā laffa
*sb/sth*_i and-REL.HUM/REL.NONHUM turn.round.PST
 laff-a-*hu*_i
 turning.round-ACC-M.3SG.POSS_i
*sb*_i and that who did just as they did, *sb*_i and that who is of their type, or *sth*_i
 and that which is of its kind

In (96), the nominal adjunct phrase is *murūr-a l-kirām* “the passing of noble people”. The idiom also contains a prepositional adjunct phrase headed by *bi-* “by” (or sometimes *‘alā* “by”). This prepositional phrase has been preposed to precede the nominal adjunct in several corpus examples. In (97), the nominal adjunct phrase is *laff-a-hu* “his turning-round”. It contains a slot for a possessive that is supplied by the context.

These nominal phrases serve the function of what is called in Arabic grammar *al-maf‘ūl-u l-muṭlaq* “the absolute object”, or *the absolute accusative* in Holes’ (2004) terms. According to Badawi *et al.* (2004: 145), this is a verbal complement that:

is mostly a verbal noun, usually cognate with its operating verb, in principle always indef[inite], and serving to emphasize the fact of the verb’s occurrence. It can also denote the number of times an act has been carried out, or the type of action and, when expanded by adj[ectival] qualification [. . .] or annexation [. . .], the additional content gives the sense of an English adverb of manner.

They also state that this verbal noun “strengthens the meaning of the action of the verb” (p. 735) and that the structure “is usually translated adverbially or periphrastically” into English (p. 451).

None of these adjunct phrases has shown any variation in its position relative to the head verb. The restriction on the position of this element in the clause may be accounted for by the fact that the absolute object does not typically precede its verb. This is the conclusion one may come to, considering the examples given for the structure in works such as Badawi *et al.* (2004) and Ryding (2005). Holes (2004) also notes that “[t]here is one special type of adverbial complement which is almost always in final position: the so-called ‘absolute accusative’ ” (p. 208). Changing this order would therefore be highly marked in text, or maybe, in some cases, odd and ungrammatical.

On the other hand, nine of the prepositional adjunct phrases have changed their positions vis-à-vis the other non-verbal idiomatic constituents in some of their corpus occurrences. These examples include the prepositional adjunct phrases in eight verbal idioms, e.g. (98) to (101), and one verb-subject idiom, i.e. in (102):

- (98) a. *ḍaraba* ‘uṣfūr-ayni *bi-ḥajar-i-n*
hit.PST sparrow-DU.ACC.INDF with-stone-GEN-INDF
he killed two birds with one stone
- b. *’aṣāba-t* *hādhihi* *l-qawmiyyat-u*
hit.target.PST-F DEM DEF-nationalism-NOM
wa-bi-ḥajar-i-n *wāḥid-i-n* ‘uṣfūr-ay
and-with-stone-GEN-INDF one-GEN-INDF sparrow-DU.ACC
lubnān-a *wa-l-falastīniyyīna* (Al-Hayat, 1997)
Lebanon-GEN and-DEF-palestinians

- lit. this nationalism hit, and with one stone, the two sparrows of Lebanon and the Palestinians
- (99) a. ghariqa ḥattā 'uthun-ay-*hi* fī *shay'*
sink.PST up.to ear-DU.GEN-M.3SG.POSS in *sth*
he became deeply involved in *sth*
- b. ghariq-tu fī qaḍiyyat-i l-sharq-i l-'awsaṭ-i
sink.PST-1SG in problem-GEN DEF-East-GEN DEF-middle-GEN
ḥattā 'uthun-ay-ya (Al-Ahram, 1999)
up.to ear-DU.GEN-1SG.POSS
- lit. I sank in the problem of the Middle East up to my [two] ears
- (100) a. *shakhṣ/shay'*_i ḥafara qabr-a *shay'*_j bi-yad-i-*hi*
*sb/sth*_i dig.PST grave-ACC *sth*_j with-hand-GEN-M.3SG.POSS
*sb/sth*_i put an end to *sth*_j
- b. bada'-nā nasma'-u 'aṣwāt-a-n ta'lū hunā
start.PST-1PL hear.1PL-IND voice.PL-ACC-INDF rise.PRS.F.3PL here
wa-hunāka taḥfur-u bi-yad-ay-hā qabr-a
and-there dig.PRS-IND with-hand-DU.GEN-F.3PL.POSS grave-ACC
l-'adab-i lladhi ltazama qaḍāyā
DEF-literature-GEN REL commit.PST cause.PL
'ummat-i-*hi* (Al-Thawra, after 2000)
nation-GEN-M.3SG.POSS
- lit. we started to hear voices rising here and there, digging with their own hands the grave of the literature that committed to the causes of its nation
- (101) a. qalaba zahr-a l-mijann-i li-*shakhṣ/shay'*
turn.around.PST back-ACC DEF-shield-GEN to-*sb/sth*
he became hostile to *sb/sth*
- b. qalab-ū la-hu zahr-a l-mijann-i (Al-Ahram, 1999)
turn.around.PST-3PL to-M.3SG back-ACC DEF-shield-GEN
lit. they turned around towards him the back of the shield
they became hostile to him
- (102) a. *lā* taqūm-u li-*shakhṣ/shay'* qā'imat-u-n
NEG stand.up.PRS.F-IND for-*sb/sth* pillar.F-NOM-INDF
sb/sth is completely ruined e.g. socially or professionally
- b. lawla l-naqd-u wa-l-rūḥ-u
if.it.were.not.for DEF-criticism-NOM and-DEF-spirit-NOM
l-naqdiyyat-u la-mā qāma-t qā'imat-u-n
DEF-critical-NOM EMPH-NEG stand.up.PST-F pillar-NOM-INDF
li-l-ḥaḍārat-i l-gharbiyyat-i mundhu l-qarn-i
for-DEF-civilization-GEN DEF-Western-GEN since DEF-century-GEN
l-sādis-i 'ashar ḥatta l-yawm-i (Al-Hayat, 1997)
DEF-sixth-GEN ten until DEF-today-GEN
- lit. and so if it were not for criticism and the critical spirit, none of the pillars of the Western civilization would have stood up since the sixteenth century until today

In (98b), the adjunct has been preposed and used as a parenthetical phrase between the verb and its object. The phrase has been marked by two commas in its original context, and *wa-* “and” has been used to introduce it. This change lays emphasis on the semantics of the prepositional phrase.

In (99b), the context shows that the idiom is used as a translation of part of the response of the former American president Bill Clinton to an allegation at a press conference that he had neglected the problem of the Middle East. As other elements in the context also show, preposing the adjunct phrase here serves his aim to counter this allegation by emphasizing his *deep* involvement in finding a solution to this problem. As noted earlier, placing emphasis on a certain element may be needed if the speaker is being challenged and thinks that they need to emphasize a specific element in response to the challenge.

The idioms in these two examples are metaphoric and isomorphic. Here, the adjuncts bear parts of their idiomatic meanings. Therefore, these meaning parts may become subject to emphasis by preposing them in appropriate contexts.

In examples (100) to (102), the end-weight principle is in operation. In the isomorphic idiom in (100), the object has been modified by a relatively long noun phrase. This modification has led the heavy object phrase to occur at the end of the clause. In (101), the object phrase *zahr-u l-mijann* “the back of the shield” is lexically heavier than the adjunct phrase *la-hu* “to him” which realizes here as a preposition and an object enclitic form. This seems to determine their order relative to each other. Finally, in (102), since the noun phrase that functions as the object of the preposition is relatively heavy, the adjunct has followed the idiomatic subject.

Although isomorphism seems to account (partially) for the occurrence of this phenomenon in examples (98) to (100), other isomorphic idioms have shown no variation in the position of their adjuncts. The possible interference of some stylistic factors cannot be ruled out. That is, preposing the prepositional phrase is normally conspicuous in text, and, therefore, may not be expected to occur frequently in general usage. Badawi *et al.* (2004) point out that “[i]nversion [. . . of constituent order] with verbal sentences is rare and mostly restricted to preposing of direct and other objects or phrases in highly rhetorical contexts” (p. 346). It may be added that these idioms have not occurred in structures where preposing the adjunct is needed for syntactic reasons.

5.4.3.2.3 COORDINATE REVERSAL

Four idioms in the data feature coordination: two verbal and two nominal idioms. As it has become evident from the analysis, all these allow for coordinate reversal:

- (103) a. 'arghā wa-'azbada
foam.PST and-froth.PST
or
b. 'azbada wa-'arghā
froth.PST and-foam.PST
he fumed with rage

- (104) a. *saraḥa* *wa-mariḥa*
 move.freely.PST and-enjoy.oneself.PST
 or
 b. *mariḥa* *wa-saraḥa*
 enjoy.oneself.PST move.freely.PST
 he did as he liked
- (105) a. *allutayyā* *wa-llatī*
 REL.F.SG.DIMIN and-REL.F.SG
 or
 b. *allatī* *wa-llutayyā*
 REL.F.SG and-REL.F.SG.DIMIN
 much ado, or lengthy difficult discussions
- (106) a. *shadd-u-n* *wa-jadhb-u-n*
 tugging-NOM-INDF and-pulling-NOM-INDF
 or
 b. *jadhb-u-n* *wa-shadd-u-n*
 pulling-NOM-INDF and-tugging-NOM-INDF
 great tension or struggle (e.g. in a strained political or social state)

In all of these examples, switching the coordinate order does not seem to trigger any semantic or discursive changes. A number of factors seem to account for the occurrence of this phenomenon and for its lack of any semantic and discursive effects. First, the conjunction *wa-* “and” does not entail any ranked or sequential order of the elements that it coordinates (see Badawi *et al.* 2004: 541ff). Therefore, even in semantically compositional usage, reversing the order of the coordinates in the case of *wa-* does not usually trigger any major differences in meaning or function.

In addition, all these idioms are non-isomorphic. The idiomatic meanings cannot be devolved over the formal constituents of the expressions. Moreover, the idiom in (105) is both lexically and syntactically idiosyncratic. It has no literal meaning for native speakers of Arabic today. Finally, even when the example in (106) is used as an isomorphic idiom (see section 3.3), reversing the order of its coordinates does not bring about any noticeable meaning changes, since these are synonymous.

5.4.3.3 *Embedding*

Embedding takes place when the components of the idiom are distributed over two clauses: a matrix clause and its subordinate one. Eighteen idioms exhibit this type of syntactic behavior: three verb-subject idioms, eleven verbal idioms and four nominal idioms. The embedded components have appeared within three different types of dependent clauses, to wit: relative clauses, complement clauses and adjunct clauses. In the following three sections, examples of these types are provided. Then, some explanations for the occurrence of this syntactic behavior in the data are discussed.

As the examples below demonstrate, whether or not the dependent clause contains a subordinator or a relative pronoun is contingent upon one or more of the following factors: the kind of embedding involved, the complementation pattern of the head of the matrix clause (e.g. the head verb might or might not require a complementizer), the grammatical characteristics of the head noun modified by the relative clause (e.g. it could be definite or indefinite), and the grammatical function of the subordinate clause itself (for an account of Arabic relative clauses in particular, see Badawi *et al.* 2004: 489ff.).

Table 5.7 contains information on the distribution of the idioms that show embedding over its three types in the data. Since some idioms exhibit more than one of these types in their corpus examples, the sum of the numbers incorporated here exceeds the total number of idioms that show embedding.

5.4.3.3.1 RELATIVE CLAUSES

The two verb-subject idioms that show relativization are:

- (107) a. 'afala najm-u *shakhṣ/shay'*
 set.PST star-NOM *sb/sth*
 the fame or glory of *sb/sth* ended
- b. 'āsha fī 'ajwā'-i-hā najm-a-n lā
 live.PST in atmosphere.PL-GEN-POSS star-ACC-INDF NEG
 ya'ful-u (Al-Hayat, 1996)
 set.PRS-IND
 lit. he lived in its atmosphere [as] a star [that] never sets
- (108) a. dāra-t raḥā *shay'*
 spin.PST-F quern *sth*
sth continued in a violent way
- b. al-'insān-u l-maḥūn-u fī raḥā tadūr-u
 DEF-human.being-NOM DEF-grind.PPTCP-NOM in quern spin.PRS-IND
 bayna tumūḥāt-i-hi wa-'arāqīl-i
 between ambition.PL-GEN-M.3SG.POSS and-obstacle.PL-GEN
 ṭarīq-i-hi (Al-Hayat, 1997)
 way-GEN-M.3SG.POSS
 lit. the human being who is ground in a quern [that] spins, between his
 ambitions and the obstacles of his way

Table 5.7 Distribution of the relevant idioms over the types of embedding

	<i>Relative clauses</i>	<i>Complement clauses</i>	<i>Adjunct clauses</i>
Verb-subject idioms	2	2	–
Verbal idioms	10	–	1
Nominal idioms	1	1	2

In these examples, there is no relative pronoun in the sentence structure, as the head nouns are indefinite.

Examples of verbal idioms that exhibit this type of embedding include:

- (109) a. ḥafara qabr-a shakhṣ/shay' bi-yad-i-hi
dig.PST grave-ACC sb/sth with-hand-GEN-M.3SG.POSS
he put an end to e.g. the success or career of *sb*, or he put an end to *sth*
- b. qabr-u-n li-l-ḡalām-i taḥfur-u-hu yad-u
grave.M-NOM-INDF for-DEF-darkness-GEN dig-IND-M.3SG hand-NOM
l-ḡaw'-i (Al-Hayat, 1997)
DEF-light-GEN
lit. a grave for darkness [that] the hand of light digs
- (110) a. rafa'a rāyat-a shay'
raise.PST flag-ACC sth
he supported *sth*
- b. wa-suqūṭ-u kull-i l-rāyāt-i llatī rafa'a-ha
and-downfall-NOM all-GEN DEF-flag.F.PL-GEN REL raise.PST-F.3PL
l-ḥukm-u l-ḥāliyy-u (Al-Hayat, 1996)
DEF-regime-NOM DEF-current-NOM
lit. and the downfall of all the flags that the current regime raised
and the failure of all the ideas that the current regime supported
- (111) a. yanfukh-u fī qirbat-i-n maqtū'-at-i-n
blow-IND in bagpipe-GEN-INDF cut.PPTCP-F-GEN-INDF
he is trying in vain to convince other people to take certain actions to
improve a situation
- b. fa-l-wāḡiḡ-u 'anna l-qirbat-a llatī
and.so-DEF-clear-NOM COPM DEF-bagpipe-ACC REL
nanfukh-u fī-hā laysa-t maqtū'at-a-n faḡasb
blow.1PL.PRS-IND into-F.3SG NEG-F.3SG cut.PPTCP-ACC-INDF only
bal 'inna-hā mutahattikat-u-n (Al-Ahram, 1999)
but indeed-F.3SG rip.apart.APTCP-NOM-INDF
lit. and so the clear thing is that the bagpipe that we are blowing into is
not just cut off, but it is ripped up
- (112) a. waḡa'a shakhṣ/shay' 'ala l-raff-i
put.PST sb/sth on DEF-shelf-GEN
he made *sb/sth* no longer useful or desirable
- b. wa-sta'ād-ū taḡālīd-a l-'ajḡād-i
and-regain.PST-3PL tradition.PL-ACC DEF-grandfather.PL-GEN
min-a l-raff-i lladhī waḡa'a-hā 'alayhi
from DEF-shelf-GEN REL put.M-F.3PL on.M.3SG
l-'ābā'-u (Al-Hayat, 1996)
DEF-father.PL-NOM
lit. and they got back the traditions of their grandfathers from the shelf
on which their fathers put them

and they readopted their grandfathers' traditions after these were ignored by their fathers

- (113) a. *dassa* *l-summ-a* *fi* *l-‘asal-i*
 put.secretly.PST DEF-poison-ACC in DEF-honey-GEN
 he deceptively promoted something bad (e.g. a way of thinking or behaving) by presenting it in an advantageous or attractive way
- b. *wa-sumūm-u-hum-u* *llatī* *yudhīb-ū-na-hā*
 and-poison.PL-NOM-M.3PL.POSS REL dissolve.PRS-3PL-IND-F.3PL
fi *l-‘asal-i* (Al-Ahram, 1999)
 in DEF-honey-GEN
 lit. and their poisons which they dissolve in the honey

In (109), there is no relative pronoun in the sentence structure, as the head noun is indefinite. On the other hand, the examples in (110) to (113) feature relative pronouns because the heads of the relative clauses in these cases are definite.

Finally, the nominal idiom that shows relativization has the structure N + Adj:

- (114) a. *‘ustuwānat-u-n* *mashrūkhāt-u-n*
 record-NOM-INDF crack.PPTCP-NOM-INDF
 repetitious boring talk about a particular topic
- b. *fa-hādhīhi* *‘ustuwānat-u-n* *‘aṣḥāḥa-t*
 and.so-DEM record-NOM-INDF become.PST-F
mashrūkhāt-a-n (Al-Hayat, 1997)
 crack.PPTCP-ACC-INDF
 lit. and so this is a record [that] has become cracked

There is no relative pronoun in the Arabic example, as the head noun is indefinite.

5.4.3.3.2 COMPLEMENT CLAUSES

In three examples, the embedded components are used within a dependent clause that functions as a complement of the head of the matrix clause. In these cases, the presence of this dependent clause is required by the complementation pattern of the head verb and is necessary for the syntactic well-formedness of the sentence.

The complement clauses in these examples perform one of two grammatical functions: the subject complement or the object complement. In the former, the dependent clause describes the subject of the main clause. In the latter, the dependent clause describes the object of the main clause. An example of the subject complement function is:

- (115) a. *‘afala* *najm-u* *shakḥṣ/shay’*
 set.PST star-NOM sb/sth
 the glory or fame of *sb/sth* ended

- b. bada'a najm-u faransā ya'ful-u tadrijiyyan fi
 begin.PST star-NOM France set-IND gradually in
 miṣr-a (Al-Hayat, 1996)
 Egypt-GEN
 lit. the star of France began [to] set gradually in Egypt

The head verb *bada'a* “to begin” does not require the presence of a complementizer.

The following two examples of the object complement function involve the use of the causative verb *ja'ala* “to make” as the head of the matrix clause:

- (116) a. al-yad-u l-'ulyā
 DEF-hand.F-NOM DEF-upper.F
 the upper hand
 b. wa-taj'al-u yad-a-hā hiya
 and-make.PRS.F.3SG-IND hand-ACC-F.3SG.POSS F.3SG
 l-'ulyā (Al-Hayat, 1996)
 DEF-upper
 lit. and it makes its hand the upper [one]
 and it gives it the upper hand
- (117) a. tajammada l-dam-u fi 'urūq-i sb
 freeze.PST DEF-blood-NOM in vein.PL-GEN sb
 the blood froze in *sb*'s veins
 b. 'amjād-u-n ka-hādhihi taj'al-u l-dam-a
 glory.PL-NOM-INDF lik-DEM make.PRS-IND DEF-blood-ACC
 yajmud-u fi l-'urūq-i (LDC Arabic Corpus)
 freeze.PRS-IND in DEF-vein.PL-GEN
 lit. glories like these make the blood freeze in the veins
 glories like these freeze the blood in the veins

5.4.3.3.3 ADJUNCT CLAUSES

In this type of embedding, the presence of the subordinate clauses is not required by the complementation pattern of the head of the matrix clause. Therefore, they could, in principle, be removed without affecting the grammatical well-formedness of the sentence.

This phenomenon occurs in three idioms. The dependent clauses in these function as adjuncts of reason, time, and purpose, as shown in the examples below, respectively:

- (118) a. dassa l-summ-a fi l-'asal-i
 put.secretly.PST DEF-poison-ACC in DEF-honey-GEN
 he deceptively promoted something bad (e.g. a way of thinking) by presenting it in an advantageous or attractive way

These form-meaning connections appear to enable the speaker or writer to reshape the syntactic structure and semantic representation of the idiom while simultaneously preserving its conventionalized non-literal meaning. The idiomatic meaning is preserved by keeping the original figurative image intact across the variations, i.e. by maintaining the deep logical relationships between the idiomatic components as they are found in the base form.

As to relativization in particular, *idiom-internal relativization* in his terms, Langlotz (2006: 258) points out that it:

is a constructional adaptation that changes the relative prominence of the profiled conceptual substructures. More specifically, it places the idiomatic NPs in focal position: the NP is fronted and topicalised [. . .]. Idiom-internal relativisation is only compatible with and functional for analysable idiomatic expressions. [. . .] For non-analysable idioms (both motivated and opaque ones) the construal of the literal scene according to this schema is not compatible with the idiomatic meaning. Due to this constraint, they cannot be manipulated accordingly.

The unanalyzability of non-isomorphic idioms appears to impose restrictions on their potential for this type of syntactic malleability. This said, the example in (117) is mainly a metonymy-based, non-isomorphic idiom that has undergone an embedding process to convey the meaning of causation. The transparency of the underlying figuration of the idiom and the general nature of the meaning of causation may account for this use.

Although isomorphism may account for the occurrence of most of the examples of this phenomenon, many isomorphic idioms have not been used in embedded constructions. Such variations are highly marked in their texts. This markedness could be ascribed to the fact that they generally call for more analytical processing effort. This need for more processing effort stems from the fact that in embedding the expression departs from its conventional structure and cannot, therefore, be cognitively processed by simply resorting to a pre-existing representation of it in the mental lexicon of the native speaker. Such highly marked uses may not be expected to occur frequently in general language use, which could provide a plausible explanation for the infrequency of this phenomenon in the data.

It may be appropriate here to point to Carter's (2004) discussion of what he terms *pattern reforming*. In this regard, he notes that some uses are more likely to be noticed because they are more overt. As he puts it, an example of pattern reforming would draw attention to itself:

because it involves a deliberate play on or with words and it is overt because it can involve a break with an expected pattern. For example, a second meaning of a word is picked up on or a fixed expression is displaced or a newly coined or invented word or phrase is produced.

(Carter 2004: 102)

5.4.3.4 Predicativization

In predicativization, an idiom that does not involve a subject-predicate relationship in its base form undergoes syntactic changes that lead its components to form (parts of) the subject and predicate of the same clause.

A number of points are in order. First, the discussion of this phenomenon obviously does not relate to verb-subject idioms, as they involve a subject-predicate relationship in their base forms. Second, the examples discussed here do not include cases of passivization wherein the object of a verbal idiom becomes the subject of a passivized version of its verb (see section 5.4.2). Finally, this section is not concerned with examples of perspective-adaptation where e.g. the object of a causative verb becomes the subject of a reflexive one (that is based on the same root) (see section 5.3.2).

Predicativization has occurred in three nominal idioms and five verbal ones. Not surprisingly, all the nominal idioms showing this pattern of syntactic behavior have the structure: head noun + adjectival modifier. This structure can be readily altered into the structure: a subject + a predicate, mainly by using the head noun in the subject position and using the modifying adjective predicatively rather than attributively. Some examples of this phenomenon are:

- (121) a. al-yad-u l-‘ulyā
 DEF-hand.F-NOM DEF-upper.F
 the upper hand
- b. naḥnu ‘alā thiqat-i-n bi-’anna-nā
 IPL on confidence-GEN-INDF with-COMP-1PL
 sa-nuḥaqqiq-u dhālika wa-’inna yad-a-nā
 FUT-achieve.IPL-IND DEM and-indeed hand.F-ACC-1PL.POSS
 sa-takūn-u hiya l-‘ulyā (Al-Tajdid, 2002)
 FUT-be-IND F.3SG DEF-upper
 lit. we are confident that we will achieve that, and, indeed, our hand
 will be the upper [one]
- (122) a. ’uṣṭuwānat-u-n mashrūkh-at-u-n
 record.F-NOM-INDF crack.PPTCP-F-NOM-INDF
 repetitious boring talk about a particular topic
- b. al-’uṣṭuwānāt-u l-’amrīkiyy-at-u laysa-t
 DEF-record.F.PL-NOM DEF-American-F-NOM NEG-F
 mashrūkh-at-a-n (Al-Thawra, after 2000)
 crack.PPTCP-F-ACC-INDF
 lit. the American records are not cracked

These examples involve changes only in the grammatical relationships between the idiomatic parts.

On the other hand, in addition to such changes, predicativizing verbal idioms might involve alternations to the morphological properties of the idiomatic words. The following are three examples of predicativization in verbal idioms. While the

“two eyes” is interpreted metaphorically in the source domain by attributing a human quality to it, i.e. the quality that a human being may or may not be believed as a source of information. In the target domain, i.e. the domain of the idiomatic meaning, on the other hand, the eyes, as the organ of vision, stand in a metonymic relationship with the what-is-seen part of the idiomatic reading.

In the predicatized version, however, the eyes seem to stand metonymically for the entity referred to by the possessive pronoun, i.e. the person who is surprised, or perhaps for their cognitive faculty. This semantic shift may be based on the fact that the eyes constitute the part of the body which is involved instrumentally in seeing and, in turn, perceiving the world.

An important general conclusion that may be drawn from the foregoing analysis and discussion is that the degree of the syntactic flexibility of Arabic idioms is much higher than that suggested by some views in contemporary Arabic linguistics. Indeed, idioms that allow for predicatization, changes in their constituent order, and/or embedding call into question the conclusion that Attia (2006) comes to:

As passivisation in Arabic is not made by configurational restructuring of the sentence, but rather by morphological inflection of verbs, we can say that Arabic shows only one instance of syntactic flexibility in MWEs [i.e. multiword expressions], that is allowing intervening elements (p. 95)

where his characterization and examples of multiword expressions clearly subsume idioms as understood in the context of the present study.

6 Conclusions

The following sections address four main points: an overview of the major findings of the study and their interpretations; main areas of application in which the findings would be useful; limitations of the study; and ideas for future research on Arabic idioms.

6.1 Major findings and their interpretations

This study has investigated the linguistic behavior of a sample of Arabic idioms chiefly on the basis of their occurrences in the *All Newspapers* section of *Arabi-corpus*. It regards *idiom* as: a multiword unit that has a syntactic function within the clause and has a figurative meaning in terms of the whole or a unitary meaning that cannot be derived from the meanings of its individual components.

The study examines the data with regard to their semantic structure, discursive behavior, and lexical and grammatical properties. The major findings relevant to each of these aspects and their plausible interpretations are provided below.

6.1.1 *The semantic structure of idioms*

A general finding that should be pointed to here is that corpus investigations have shown that idioms are sometimes used in meanings that are not recorded for them in available Arabic dictionaries.

The study has examined two major points with respect to the semantic structure of Arabic idioms: the patterns of semantic extension that underlie them and how they can be classified in regard to the notion of isomorphism.

6.1.1.1 *Patterns of semantic extension in idioms*

In section 3.2, the study has found, first, that the main patterns of semantic extension that are involved in the production of Arabic idioms are metaphor, metonymy, interaction of metaphor and metonymy, semantic extension based on conventional knowledge, hyperbole, and emblemizing.

Second, some correlations have been identified between the patterns of figuration and the syntactic classes of idioms. In particular, many prepositional idioms

contain a metonymic element. This correlation is explainable by the range of the meanings that Arabic prepositions typically convey, e.g. various associations in time or space or different relationships of causality, instrumentality, and accompaniment. Since metonymy is, by definition, concerned with meaning extensions that are based on such relationships, when prepositional phrases are used figuratively, this use often involves a metonymy.

Third, four types of interaction between metaphor and metonymy have been identified in the data. Three of these, i.e. metaphor from metonymy, metonymy within metaphor, and metaphor within metonymy, are discussed in Goossens (2003). The fourth type involves a mapping between an entailment of the literal reading and the idiomatic meaning, and it is not discussed in Goossens' work.

Fourth, the metaphor from metonymy type is by far the most frequent among these patterns of interaction in the data. The high frequency of this pattern may be partially ascribed to its experiential basis. The other infrequent patterns of metaphor-metonymy interaction do not in general have the degree of embedding in human everyday experience that this pattern has.

Additionally, as far as the metonymy within metaphor pattern is concerned, the condition to incorporate an element that functions both in the source and target domains could be regarded as a restriction on the productivity of this type. However, in this regard, it is stressed that in order to better understand this phenomenon, future work needs to be conducted on different sets of data with some effort to categorize idioms in semantically coherent groups and with close attention to the question of how distinct the target domains are from the source domains. Finally, the infrequency of the metaphor within metonymy type may be explained by “[t]he fact that if we embed a metaphor into a metonymy, it tends to ‘metaphorise’ the whole expression” (Goossens 2003: 367).

Finally, idioms that are based on the metaphor from metonymy pattern vary with regard to their ability to be used as pure metonymies. This may be due to linguistic and/or socio-cultural factors. For example, in principle, clearly motivated idioms which are relevant to the present-day world of native speakers can be used as metonymies. On the other hand, literally non-compositional idioms may not be used in this way. The restriction on the purely metonymic uses could also occur because of socio-cultural changes in the speech community. For instance, an idiom that belongs to this type may continue to be used, while the situation with which it is metonymically associated may cease to occur in the real world today. Also, the metonymic basis of a motivated idiom might be irrelevant to everyday life of native speakers because the expression is originally a result of a borrowing process that has taken place due to some form of language contact.

6.1.1.2 Idioms and isomorphism

In section 3.3, the study has shown, first, that a binary classification of idioms into isomorphic and non-isomorphic idioms is not sufficient. A third class is needed to accommodate examples that can be used in both ways depending on

context. The categorization is based on how native speakers of Arabic perceive idioms as indicated by the corpus data. The major reason for the need for this third category is that (different) speakers' perceptions of the same idiom may vary. Sometimes, they seem to perceive the idiom as having a non-literal meaning whose internal semantic structure matches that of its literal reading. At other times, the same idiom is used in a meaning that is based on an abstraction from an entailment of its literal reading. In this case, the internal make-up of the idiomatic meaning cannot be mapped onto that of the literal interpretation.

Second, while metaphorical idioms are usually isomorphic, metonymy-based idioms are typically not. These correlations have been linked to insights from some studies on metaphor and metonymy. Metaphorical idioms are typically based on noticing a set of similarities between two concepts. This usually leads to their isomorphism, as it is often possible to establish the connections between parts of the idiomatic meaning and parts of the constituent structure encoding the literal meaning. However, the study shows that some metaphorical idioms are not isomorphic due to their syntactic structure.

On the other hand, the presence of a metonymy in the semantic structure of the idiom usually renders it non-isomorphic. This may be explained by the nature of the relationships that link the two concepts in the case of metonymy. These relationships are not based on similarity but rather on connections of other types such as those between the part and the whole, the cause and the effect, and the producer and the product, to name just a few. On the whole, unlike the similarity relationship, such connections do not enable the language user to establish one-to-one links between identifiable parts of the idiomatic meaning and elements of the constituent structure of the relevant idioms.

However, it has become evident that an idiom that involves a metonymy may be isomorphic in two cases: (a) when the metonymy operates at the individual-constituent level rather than at the level of the whole expression, or (b) when the metonymy operates at the level of the whole expression but the idiom includes a constituent that is motivated by a separate pattern of semantic extension such as emblemizing.

6.1.2 *The discursive behavior of idioms*

In Chapter 4, the study has addressed two main issues with regard to the discursive behavior of idioms: the discursive functions that they perform and the ways in which they contribute to the cohesion of their texts.

6.1.2.1 The discursive functions of idioms

In section 4.1, the study shows that Arabic idioms perform different types of functions, including informational, evaluative, modalizing, and organizational functions. However, the organizational function is relatively infrequent in the data, while the informational, evaluative, and modalizing functions are common. Confining what is considered as *idiom* to multiword units that occur within the

clause may be a credible explanation for this dissimilarity. Previous research on multiword units demonstrates that the organizational function is often carried out by elements that occur outside the clause structure and that, on the other hand, the informational and evaluative functions often realize through elements that occur as parts of the clause (Fernando 1996).

The figurativity of many idioms in the data could constitute an additional explanation for the low incidence of the organizational function (see Moon 1998).

The high frequency of the informational function can also be ascribed to the fact that giving information is an essential function of language (Moon 1998). Furthermore, the property of figurativity in idioms may explain why the evaluative function is frequent in the data. As instances of figurative language, many idioms bring to mind scenes that may trigger a number of implicatures. The scene expressed in the literal reading may be culturally regarded approvingly or disapprovingly and this may provide the basis for the evaluative content in these expressions.

The high incidence of the modalizing function, which often manifests in situating an action or quality at the extreme end of a scale between two opposites, may be explained by the fact that idioms are not often used to refer to regular phenomena or moderate qualities. Rather, as previous research shows, metaphors in general and idioms are used to convey intense versions of the meanings of any apparently equal literal paraphrases.

The study has identified many instances of multiple functioning wherein an idiom performs more than one discursive role simultaneously. Considered together, the properties discussed above in accounting for the pervasiveness of the informational, evaluative, and modalizing functions may help explain this phenomenon.

The study has also identified some correlations between the discursive functions and the syntactic classes of idioms. For example, the organizational function in the data is carried out by prepositional and adverbial idioms. This could be explained by the fact that prepositions and adverbs usually express different types of spatial, temporal, and logical relationships.

A salient point in the data is that negatively evaluative idioms outnumber positively evaluative ones. This phenomenon may be explained by the property of *indirectness* in idioms. Since idioms develop as indirect ways to communicate their meanings, this often makes them an appropriate choice when indirectness is needed to talk about taboo topics and/or express negative evaluations.

In general, language users need tactful ways to express negative evaluations more than they do to express positive ones. This could provide a motivation for formularizing more indirect negatively evaluating expressions than indirect positively evaluating ones. Since the use of formulaic language is characterized by the need for less cognitive processing effort (Wray 2002), formularizing many negatively evaluative non-literal expressions could then be regarded as an attempt to maximize the communicative advantages that these expressions equip language users with, particularly their ability to allude to shared values and to create and support rapport between speakers or writers and hearers or readers.

However, the study notes that even though the evaluative content in idioms is conveyed indirectly, this does not mean that it is necessarily conveyed politely. Two reasons may be provided for this. First, conventionalizing the form-meaning relationship makes the idiom more “direct” in expressing its non-literal meaning. In this regard, Brown and Levinson’s (1987) notion of *conventionalized indirectness* is of great relevance. Second, the implicatures that are often evoked by the literal senses of many idioms, especially when they are elaborated in context, may intensify their evaluative content.

Two major sources of explanations have been identified to account for the particular orientation of the evaluation in idioms. The first is the characteristics of the images underlying them, perhaps including how the literal sense is interpreted within its etymological socio-cultural context. The other source is the values of the speech community in their dynamic, ever-changing nature.

Some evaluative idioms in the data can undergo changes in the typical orientation or intensity of their evaluative content. This has only been attested in the case of clearly motivated examples.

In general, the clarity of the motivation of the idiom has been a powerful explanatory notion in the context of discussing the discursive functions of idioms. Since this part of the study is mainly concerned with how idioms contribute to the content of their texts, this notion, with its focus on the meaning-meaning relationship, is a good candidate for this task.

6.1.2.2 Idioms and cohesion

In section 4.2, the study has identified six major ways in which Arabic idioms contribute to the cohesion of their texts. These include the use of determiners and demonstratives, signaling temporal, spatial, or logical relations through organizational idioms, the use of some types of syntactic modifiers, the extension of the figurative image underlying the idiom, contextual transformations, and some of the morphological properties of the idiomatic components. The isomorphism of idioms has played the major part in explaining the ability of idioms to contribute to the cohesion of their texts through most of these ways or devices.

6.1.3 The lexical and grammatical behavior of idioms

In Chapter 5, the study has investigated the following points with regard to the formal behavior of idioms: lexical variation, changes in lexicogrammatical complexity, perspective-adaptation, inflectability, the active and passive voice, syntactic types and structures, constituent order, embedding, and predicatization.

It has become evident that Arabic idioms show a degree of formal variation that is higher than what is proposed in some contemporary accounts of the phenomenon (e.g. in Attia 2006). In particular, many examples of lexical substitutions, changes in constituent order, embedding, and predicatization have been attested in the corpus data. Once more, the isomorphism of many idioms has been the most

powerful explanatory notion in this regard. However, the clarity of the motivation has played an important role in this respect too.

On the other hand, several explanations have been proposed for the restrictions on the formal variation of idioms. First, the incompatibility between the grammatical meaning of the type of formal variation and the idiomatic meaning is an important reason for such restrictions. In this regard, one may point to some verbal idioms that do not occur in the past tense because the meaning of completion expressed by this form in Arabic may contradict their idiomatic meanings. Second, the syntactic structure of some isomorphic idioms seems to restrict their formal variation potential.

Third, some sociolinguistic factors related to the present state of diglossia in Arabic could also account for some relevant phenomena in the data, particularly with regard to the choice of the negatives and in turn the form of the verbs that follow them. Fourth, both the grammatical environments in which the idiom occurs and the characteristics of the image underlying it may account for the restrictions on the inflection of some of its nominal components. Fifth, the discursive function of the idiom could restrict its formal variability. For instance, verbal idioms that convey negative evaluations of their subjects tend not to occur in the first person form. Finally, some stylistic factors may be at work too. That is, the infrequency of some syntactic variations might be explained by the fact that they are highly marked in text and therefore may not be expected to occur frequently in general language usage.

6.2 Contexts of use and applications

The findings of this study may be of interest to several domains of application, including lexicography, Arabic language learning and teaching, natural language processing, and translation.

The analyses of individual examples may be a good source of information that could be incorporated in the entries devoted to idioms in dictionaries. Currently, these entries focus mainly on definitions. Depending on the target users of the dictionary, different types of information and different degrees of detail may be adopted. For example, information on the variation potential of idioms is of particular significance in dictionaries designed for learners of Arabic. Also, information on any evaluative function that the idiom has is crucial to ensure native-like encoding and decoding.

Likewise, authors of Arabic textbooks may adopt an equal level of detail so as to help the learner develop a fuller understanding of the expression being presented. Since some types of information on the uses of idioms are not accessible to intuition, the corpus analyses provided here may equip the authors of such books and Arabic teachers with some of the information needed in order to bring the learning outcomes closer to native-like proficiency and to make the contents of Arabic textbooks and classes more representative of the state of MSA in the real world.

Details of the formal variations that many idioms allow are crucial in the context of natural language processing. Failing to encode such information appropriately

in the system may make it impossible to recognize a particular variant of an idiom as such. This in turn could lead to some major inaccuracies in the intended outcome of the whole process, be it a translation or a grammatical account of a text, for instance.

In the context of human translation, knowledge of any evaluative function of the idiom is essential for selecting it or not in a particular context. It is also crucial in deciding on its appropriate equivalents in the target language. This is relevant even in the cases where Arabic idioms are originally loan translations from English, for instance, since the Arabic expression could develop connotations in its contexts of use that are not typical of its original English equivalent.

6.3 Limitations of the study

No corpus can include every single aspect of the language under examination. Therefore, it is likely that some relevant phenomena that do not occur in *Arabic corpus* do occur outside it. Also, the fact that it mainly consists of newspaper texts could undermine some relevant phenomena that occur in other styles. Moreover, the complications involved in searching *Arabic corpus* could sometimes make it difficult to retrieve all the relevant examples from its texts. Although the study has attempted to overcome these limitations e.g. by using additional sources of data, this attempt has been limited.

In some cases, these limitations have led to a degree of subjectivity. In this regard, one may point to the difficulty of establishing the precise discursive function(s) of some idioms that do not occur frequently in *Arabic corpus*.

It could be added here that a fuller understanding of some phenomena in the data may only be arrived at through incorporating a diachronic dimension in the investigation. However, due to the nature of the data sources available for me, no effort to do so has been made in any systematic way.

6.4 Ideas for future research

Many topics may be suggested for future research on Arabic idioms. Some of the most prominent of these are:

- a study of how conceptual metaphors participate in the production of Arabic idioms;
- a detailed investigation of syntactic modification in Arabic idioms;
- an investigation of the grammatical environments in which Arabic idioms occur, e.g. this could focus on the polarity contexts in which different Arabic idioms are (typically) used;
- an examination of how open for derivation Arabic idioms are; e.g., a study may be interested in examining the phenomenon of nominalization in verbal idioms and its functions in the case of idioms that convey negative evaluations of their subjects;

- a detailed investigation of what lexical items collocate with Arabic idioms and how different collocational patterns affect the discursive and semantic behavior of idioms;
- investigating the sources of idioms in MSA, e.g. literary works, translation and spoken dialects – this may be linked to several sociolinguistic issues in Arabic-speaking countries;
- psycholinguistic investigations of how native speakers of Arabic decode and encode idioms;
- investigating how idioms are sometimes signaled metalinguistically in text;
- a cross-corpora analysis of Arabic idioms and their English equivalents;
- developing an electronic database of Arabic idioms that includes information on their linguistic behavior, particularly their variation potential – such a database may also be developed into a bilingual work by incorporating the findings of the study suggested in the previous point.

Indeed, (corpus-based) investigations of Arabic idioms and Arabic phraseology in general that are synchronic or diachronic in nature are much needed for both theoretical and practical purposes.

Notes

2 Methodology and data

* Because many idioms in MSA are loan translations from English, sometimes my attention was drawn to the Arabic idiom only after coming across its English counterpart in my daily interactions and readings.

† *Arabicorpus* is available at <http://arabicorpus.byu.edu/>

‡ The search terms can be typed using the Arabic script too. But I have often relied on the transliterations because they are more straightforward when regular expressions are involved.

3 The semantic structure of Arabic idioms

* This term is used broadly in her work, so that it subsumes many of the phenomena that are considered as idioms in the present work.

† This is only one possible basis for this idiom. Since the word *qā'imah* can mean “an animal’s leg”, another possibility is that the idiom is based on an image of an animal that has none of its legs standing.

5 The lexical and grammatical behavior of Arabic idioms

* Also, see his use of *conspicuousness* as a criterion for wordplay (Langlotz 2006: 222). However, he admits the difficulty of setting the boundaries of this principle.

† However, rules of general logic are not always observed in language structure, and many idioms are archetypes of this. In fact, it has been said that “[i]f natural language had been designed by a logician, idioms would not exist” (Johnson-Laird 1993: vii).

‡ In a spoken version of this example, the vowel /a/ in *wa-* “and” could be dropped. Also, *fēn* represents the colloquial interrogative *فين* “where”.

§ Ryding (2005) also discusses what she calls the derivational passive in Arabic, where “a derivational verb form [. . .] is used to convey a passive, reflexive, or mediopassive sense of the action involved in the verb” (p. 657). However, she acknowledges Wright’s (1967) point that the meanings conveyed by the verb patterns involved in this type are resultative, and, therefore, are not completely identical with the concept of formal passive. She states that in English “it is sometimes necessary to render the equivalent meaning of these derived forms in the passive” (p. 657). This type of “passive” is not a concern of this section, since its most prominent examples in the data have been discussed in section 5.3.2.

¶ The only account I have found of the figuration underlying this classical idiom involves an image of a person who, when confused or regretful, starts to bite their hand, thus making their mouth “fall on” their hand. Given this description, this opaque idiom seems to have a metonymic basis.

- || A few examples of this idiom in the SV order have been found in other sources, mainly websites. However, the texts in which they occur belong to Egyptian colloquial Arabic, wherein the SV order is predominant. These examples have therefore been excluded from the study. In a relevant context, Holes (2004) states that even “in the case of ‘event-oriented’ sentences, the preference for VSCOMP is to some degree a matter of personal choice. Some journalists occasionally use SVCOMP where it is clearly an ‘event’ which is being described but for no contextually obvious reason. A contributory influence may be the writer’s dialectal background: ‘urban’ dialects, like Cairene and Damascene, which are spoken by the majority of writers and journalists, have SVCOMP as the normal order for all types of message. Another possible influence may be the typically SVCOMP order of European languages, especially English, with which many Arab writers are familiar and which have a predominant position in news dissemination and in the cultural transfer of technical information of all kinds” (p. 253). His reference to the position of English in the transfer of technical knowledge seems to be connected with the function of SVCOPM order in expository writing as he describes it in an earlier quotation.

Bibliography

- Abdel-Aziz, M. H. (2003) *luġhat-u l-šihāfat-i l-mu'āširat-i*. [The language of contemporary journalism]. Cairo: Dār Al-Fikr Al-'arabiyy.
- Afonso, S. C. P. (2008) *The family of impersonal constructions in European Portuguese: An onomasiological constructional approach*. Unpublished PhD thesis. The University of Manchester.
- Al-Sulaiti, L. (2004) *Designing and developing a corpus of contemporary Arabic*. Unpublished MSc thesis. University of Leeds.
- Attia, M. (2006) Accommodating multiword expressions in an Arabic LFG grammar. In T. Salakowski, F. Ginter, S. Pyysalo and T. Pahikkala (eds) *Advances in natural language processing. Lecture notes in Computer Science*, 4139, pp. 87–98. Berlin: Springer.
- Badawi, E. (1995) The use of Arabic in Egyptian T.V. commercials: A language simulator for the training of teachers of Arabic as a foreign language. In J. Alatis and C. Strachle (eds) *Linguistics and the education of language teachers: Ethnolinguistic, psycholinguistic, and sociolinguistic aspects*, pp. 33–39. Washington, DC: Georgetown University Press.
- Badawi, E., Carter, M. G. and Gully, A. (2004) *Modern written Arabic: A comprehensive grammar*. London: Routledge.
- Barkema, H. (1996) Idiomaticity and terminology: A multi-dimensional descriptive model. *Studia Linguistica*, 50, 125–60.
- Bennett, P. (2002) *Semantics: An introduction to non-lexical aspects of meaning*. Muenchen: LINCOM Europa.
- Breidt, E., Segond, F. and Valetto, G. (1996) Local grammars for the description of multiword lexemes and their automatic recognition in texts. In *Proceedings of 4th Conference on Computational Lexicography and Text Research, COMPLEX'96*, pp. 19–28. Budapest.
- Brinton, L. J. and Traugott, E. C. (2005) *Lexicalization and language change*. Cambridge: Cambridge University Press.
- Brown, P. and Levinson, S. C. (1987) *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Burger, H. (1998) *Phraseologie: eine Einführung am Beispiel des Deutschen*. Berlin: Erich Schmidt Verlag.
- Cameron, D. (2001) *Working with spoken discourse*. London: SAGE Publications.
- Carter, R. (2004) *Language and creativity: The art of common talk*. London: Routledge.
- Channell, J. (2001) Corpus-based analysis of evaluative lexis. In S. Hunston and G. Thompson (eds) *Evaluation in text: Authorial stance and the construction of discourse*, pp. 38–55. Oxford: Oxford University Press.

- Conrad, S. and Biber, D. (2001) Adverbial marking of stance in speech and writing. In S. Hunston and G. Thompson (eds) *Evaluation in text: Authorial stance and the construction of discourse*, pp. 56–73. Oxford: Oxford University Press.
- Cowie, A. P. (1983) Introduction. In A. P. Cowie, R. Mackin and I. R. McCaig (eds) *Oxford dictionary of current idiomatic English*. Oxford: Oxford University Press.
- (1988) Stable and creative aspects of vocabulary use. In R. Carter and M. McCarthy (eds) *Vocabulary and language teaching*, pp. 126–139. London: Longman.
- (ed.) (1998a) *Phraseology: Theory, analysis, and applications*. Oxford: Clarendon Press.
- (1998b) Introduction. In A. P. Cowie (ed.) *Phraseology: Theory, analysis, and applications*, pp. 1–20. Oxford: Clarendon Press.
- (2006) Phraseology. In K. Brown (ed.) *Encyclopaedia of language and linguistics*, pp. 579–585. Amsterdam: Elsevier.
- Croft, W. and Cruse, A. (2004) *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Deignan, A. (2005) *Metaphor and corpus linguistics*. Amsterdam: John Benjamins Publishing Company.
- Dobrovols'kij, D. and Piirainen, E. (2005) *Figurative language: cross-cultural and cross-linguistic perspectives*. Amsterdam: Elsevier.
- Fellbaum, C. (1993) The determiner in English idioms. In C. Cacciari and P. Tabossi (eds) *Idioms: Processing, structure, and interpretation*, pp. 271–296. Hillsdale, NJ: Lawrence Erlbaum Associates.
- (ed.) (2007a) *Idioms and collocations: Corpus-based linguistic and lexicographic studies*. London: Continuum.
- (2007b) Introduction. In C. Fellbaum (ed.) *Idioms and collocations: Corpus-based linguistic and lexicographic studies*, pp. 1–19. London: Continuum.
- Fernando, C. (1996) *Idioms and idiomaticity*. Oxford: Oxford University Press.
- Fillmore, C. J., Kay, P. and O'Connor, M. C. (1988) Regularity and idiomaticity in grammatical constructions: The case of “let alone”. *Language*, 64, 501–538.
- Geeraerts, D. (2003) The interaction of metaphor and metonymy in composite expressions. In R. Dirven and R. Pörings (eds) *Metaphor and metonymy in comparison and contrast*, pp. 435–465. Berlin: Mouton de Gruyter.
- Gibbs, R. (1989) Understanding and literal meaning. *Cognitive Science* 13, 243–251.
- (1994) *The poetics of mind*. Cambridge: Cambridge University Press.
- Glucksberg, S. (2001) *Understanding figurative language: From metaphor to idioms*. New York: Oxford University Press.
- Goossens, L. (2003) Metaphonymy: The interaction of metaphor and metonymy in expressions for linguistic action. In R. Dirven and R. Pörings (eds) *Metaphor and metonymy in comparison and contrast*, pp. 349–377. Berlin: Mouton de Gruyter.
- Granger, S. (2005) Pushing back the limits of phraseology. How far can we go? In C. Cosme, C. Gouverneur, F. Meunier and M. Paquot (eds) *Proceedings of the Phraseology 2005 Conference*, pp. 165–168. Louvain-la-Neuve, 13–15 October 2005. Online. Available at: <http://cecl.fltr.ucl.ac.be/Downloads/Granger%202005b.doc> (accessed 4 May 2009).
- Gross, M. (1993) Local grammars and their representation by finite automata. In M. Hoey (ed.) *Data, description, discourse: Papers on the English language in honour of John McH Sinclair*, pp. 26–38. London: Harper-Collins.
- Halliday, M. A. K. and Hasan, R. (1976) *Cohesion in English*. London: Longman.
- Hasan, R. (1994) The texture of a text. In D. Graddol and O. Boyd-Barrett (eds) *Media texts, authors and readers: A reader*, pp. 74–89. Clevedon: Multilingual Matters in association with the Open University.

- Hickey, T. (1993) Identifying formulas in first language acquisition. *Journal of Child Language*, 20, 27–41.
- Hoey, M. (1991) *Patterns of lexis in text*. Oxford: Oxford University Press.
- (2001) *Textual interaction: An introduction to written discourse analysis*. London: Routledge.
- Holes, C. (2004) *Modern Arabic: Structures, functions, and varieties*. Washington, DC: Georgetown University Press.
- Hopper, P. J. and Traugott, E. C. (2003) *Grammaticalization*. Cambridge: Cambridge University Press.
- Howarth, P. A. (1996) *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Tübingen: Niemeyer. Online. Available at: <http://www.leeds.ac.uk/languages/contact/people/howarth.html> (accessed 4 April 2007).
- (1998) Phraseology and second language proficiency. *Applied Linguistics*, 19, 24–44.
- Hümmer, C. (2007) Meaning and use: a corpus-based case study of idiomatic MWUs. In C. Fellbaum (ed.) *Idioms and collocations: Corpus-based linguistic and lexicographic studies*, pp. 138–151. London: Continuum.
- Johnson-Laird, P. N. (1993) Foreword. In C. Cacciari and P. Tabossi (eds) *Idioms: Processing, structure, and interpretation*, pp. vii–x. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Kövecses, Z. (2002) *Metaphor: A practical introduction*. Oxford: Oxford University Press.
- Krzyszowski, T. P. (1990) The axiological aspect of idealized cognitive models. In J. Tomaszczyk and B. Lewandowska-Tomaszczyk (eds) *Meaning and lexicography*, pp. 135–165. Amsterdam: John Benjamins Publishing Company.
- Lakoff, G. and Johnson, M. (2003) *Metaphors we live by*. Chicago: University of Chicago Press.
- Langlotz, A. (2005) Are constructions the basic units of grammar? – The phraseological turn in linguistic theory. In A. Hamm (ed.) *Language chunks and linguistic units*, pp. 45–63. Strasbourg: Université Marc Bloch.
- (2006) *Idiomatic creativity: A cognitive-linguistic model of idiom-representation and idiom-variation in English*. Amsterdam: John Benjamins Publishing Company.
- Louw, B. (2000) Contextual prosodic theory: Bringing semantic prosodies to life. In C. Heffer and H. Sauntson (eds) *Words in context: A tribute to John Sinclair on his retirement*. Birmingham: University of Birmingham.
- Martin, J. R. (2003) Cohesion and texture. In D. Schiffrin, D. Tannen and H. E. Hamilton (eds) *The handbook of discourse analysis*, pp. 35–53. Oxford: Blackwell Publishing.
- McArthur, M. (ed.) (1998) *Concise Oxford companion to the English language*. Oxford: Oxford University Press. Online. Available at: http://www.oxfordreference.com/views/BOOK_SEARCH.html?book=t29&authstatuscode=202 (accessed 4 February 2008).
- Minugh, D. (2007) The filling in the sandwich: Internal modification of idioms. In R. Facchinetti (ed.) *Corpus linguistics: 25 years on*, pp. 205–224. Amsterdam: Rodopi.
- Moon, R. (1998) *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Clarendon Press.
- Naciscione, A. (2001) *Phraseological units in discourse: Towards applied stylistics*. Latvia: Latvian Academy of Culture.
- Nunberg, G., Sag, I. A. and Wasow, T. (1994) Idioms. *Language*, 70, 491–538. Oxford University Press. *Oxford English Dictionary*. Oxford University Press. Online. Available at: <http://www.oed.com/>
- Palmer, F. R. (1990) *Modality and the English modals*. London: Longman.
- Parkinson, D. (2009) “Arabicorpus”. E-mail (19 June 2009).

- Perry, F. (2005) *Research in applied linguistics: Becoming a discerning consumer*. London: Lawrence Erlbaum Associates.
- Pulman, S. G. (1993) The recognition and interpretation of idioms. In C. Cacciari and P. Tabossi (eds) *Idioms: Processing, structure, and interpretation*, pp. 249–270. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Radden, G. (2003) How metonymic are metaphors? In R. Dirven and R. Pörings (eds) *Metaphor and metonymy in comparison and contrast*, pp. 93–109. Berlin: Mouton de Gruyter.
- Riemer, N. (2003) When is a metonymy no longer a metonymy? In R. Dirven and R. Pörings (eds) *Metaphor and metonymy in comparison and contrast*, pp. 379–406. Berlin: Mouton de Gruyter.
- Riehemann, S. (2001) *A constructional approach to idioms and word formation*. Unpublished PhD thesis. Stanford University.
- Ryding, K. C. (2005) *A reference grammar of modern standard Arabic*. Cambridge: Cambridge University Press.
- Saeed, J. I. (2003) *Semantics*. Oxford: Blackwell Publishing.
- Schmitt, N., Grandage, S. and Adolphs, S. (2004) Are corpus-derived recurrent clusters psycholinguistically valid? In N. Schmitt (ed.) *Formulaic sequences: Acquisition, processing, and use*, pp. 127–151. Amsterdam: John Benjamins Publishing Company.
- Sieny, M. E., Hussein, M. A. and Al-Doush, S. A. (1996) *Al-mu'jam Al-siyāqīyy li-l-ta'bīrāt Al-iṣṭilāḥīyyah*. [The contextual dictionary of conventionalized expressions]. Lebanon: Maktabat Lubnān Nāshirūn.
- Sinclair, J. (2005) Appendix: How to build a corpus. In M. Wynne (ed.) *Developing linguistic corpora: A guide to good practice*, pp. 79–83. Oxford: Oxbow Books. Online. Available at: <http://ahds.ac.uk/creating/guides/linguistic-corpora/appendix.htm> (accessed 3 April 2009).
- Sathi, K. (2007) A corpus-based analysis of adjectival modification in German idioms. In C. Fellbaum (ed.) *Idioms and collocations: Corpus-based linguistic and lexicographic studies*, pp. 81–108. London: Continuum.
- Stetkevych, J. (1970) *The modern Arabic literary language: Lexical and stylistic developments*. Chicago: University of Chicago Press.
- Strässler, J. (1982) *Idioms in English: A pragmatic analysis*. Tübingen: Gunter Narr Verlag.
- Stubbs, M. (1983) *Discourse analysis*. Oxford: Basil Blackwell.
- Teliya, V., Bragina, N., Oparina, E. and Sandomirskaya, I. (1998) Phraseology as a language of culture: Its role in the representation of a cultural mentality. In A. P. Cowie (ed.) *Phraseology: Theory, analysis, and applications*, pp. 55–75. Oxford: Clarendon Press.
- Thompson, G. and Hunston, S. (2001) Evaluation: An introduction. In S. Hunston and G. Thompson (eds) *Evaluation in text: Authorial stance and the construction of discourse*, pp. 1–27. Oxford: Oxford University Press.
- Tognini-Bonelli, E. (2001) *Corpus linguistics at work*. Amsterdam: John Benjamins Publishing Company.
- Torlakova, L. I. (1998) Some cultural and ethnic elements in Modern Standard Arabic idioms. In K. Dévényi and T. Iványi (eds) *Proceedings of the Arabic and Islamic sections of the 35th international congress of Asian and North African studies*, pp. 63–71. Budapest, 1–7 July 1997.
- Tovena, L. M. (1996) *Studies on Polarity Sensitivity*. Unpublished Ph.D. thesis. University of Edinburgh.

- Trask, R. L. (1993) *A dictionary of grammatical terms in linguistics*. London: Routledge.
- Villavicencio, A. and Copestake, A. (2002) On the nature of idioms. *LinGO Working Paper No. 2002-04*. Retrieved 6 May 2006 from: <http://lingo.stanford.edu/working-papers.html>
- Wehr, H. (1994) *Dictionary of modern written Arabic: Arabic-English*. Ed. J. Milton Cowan. Ithaca, NY: Spoken Language Services Inc.
- Wray, A. (2002) *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- (2008) *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.
- Wray, A. and Perkins, M. R. (2000) The functions of formulaic language: An integrated model. *Language and Communication*, 20, 1–28.
- Wright, W. (1967) *A Grammar of the Arabic language*. Cambridge: Cambridge University Press.
- Xiao, R. and McEnery, T. (2006) Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics*, 27, 103–129.

Index

- 'absolute accusative' 137
active voice 122–7
adjectival idioms 33; syntactic structures 132
adjective variation: idioms 107
adjunct clauses 144–5
adverbial idioms 34; syntactic structures 132
afala see "to set"
al-'asal see "the honey"
al-summ see "the poison"
al-tu'm see "the bait"
alam 41–2, 64
Arabic idioms 1–12, 150–7; adjectival idioms 33; adjectival idioms syntactic structures 132; adjective variation 107; adjunct clauses 144–5; adverbial idioms 34; adverbial idioms syntactic structures 132; changes in the lexicogrammatical complexity 109–11; classifying 16–17; cohesion 91–101; collecting 13–16; complement clauses 143–4; constituent order 133–4; context use and applications 155–6; coordinate reversal 139–40; corpus data 18–24; data examples and their frequency 26–36; discursive behavior 66–101; discursive functions 66–91; embedding 140–6; future research 156–7; idioms and cohesion 154; idioms and isomorphism 151–2; idioms discursive behaviour 152–4; idioms discursive functions 152–4; idioms grammatical behaviour 113–49; idioms inflectability 113–22; idioms lexical and grammatical behaviour 154–5; idioms lexicogrammatical behaviour 109–13; idioms semantic structure 150–2; idioms syntactic behaviour 127–33; idioms syntactic types and structures 127–33; isomorphism 55–61; language of Arabic newspapers 24–6; lexical and grammatical behaviour 102–49; lexical variation 104–9; limitations 156; major finding and interpretations 150; methodology and data 13–; model of analysis 9–10; modern standard arabic 3; motivation 37–9; nominal idioms 31–2; nominal idioms syntactic structures 131; notes on idiom frequencies 34–6; noun variation 106–7; nouns and adjectives inflectability 119–22; other types of multiword units and 4–7; perspective-adaptation 111–13; positioning the clause adjuncts 136–9; positioning the idiomatic subject 134–6; predicatization 147–9; preposition variation 107; prepositional idioms 32–3; prepositional idioms syntactic structures 132; relative clauses 141–3; relevant idioms distributions over embedding types 141; research questions 2–3; selecting the sample 17; semantic extension patterns 39–55; semantic extension patterns in idioms 150–1; semantic structure 37–65; significance of the study 1–2; sources of idioms in MSA 3–4; study 150–7; syntactic classes 16; systematic variation vs wordplay 102–4; using corpora in idiom studies 7–9; verb-subject idioms 27–8; verb-subject idioms syntactic structures 128; verb variation 105–6; verbal idioms 28–31; verbal idioms syntactic structures 129–30; verbs inflectability 113–19
Arabic newspapers 24–6

- Arabicorpus* 18–22, 109, 111, 117, 119, 150, 156; components other than newspapers 19; contents and structure 18–20; searching method 20–2; word count of each newspaper in 18
 ‘*aynan* see “two eyes”
- bala‘a* see “to swallow”
baqarah see “cow”
beat one’s breast 46
blended metaphorical idioms 44
blue blood 53–4
 British National Corpus (BNC) 21
- Classical Arabic (CLA) 3
 clauses 141–5; adjunct *see* adjunct clauses; complement *see* complement clauses; positioning the clause adjuncts 136–9; relative *see* relative clauses
 co-reference 91
 cohesion 91–101; contextual transformations 99–100; determiners and demonstratives 92–5; extending figurative image 97–9; idioms 154; modifiers 96–7; morphology of idiomatic components 100–1; organizational idioms 95–6
 cohesive tie 91
 compatibility 116
 complement clauses 143–4
 complex repetition 91
composite units 4
 constituent order 133–4
constituent motivation 38
 context: transformations 99–100; use and applications 155–6
 conventional knowledge 51–2
 coordinate reversal 139–40
corpus 7
 corpus-based research 10
 corpus data: additional data sources 22–3; *Arabicorpus* 18–22; preparing of 24; sources, collection procedures, and preparation 18–24
 “cow” 59
cranberry elements 38, 41
 cultural norms 85
 cultural values 85
- decomposable idioms* *see* idiomatic combinations
decomposition 56
 demonstratives 92–5
 determiners 92–5
- discourse: Arabic idioms 66–101; discursive functions of idioms 66–91; distribution of data over functional categories 68–70; distribution of idioms over discursive functions 69; evaluative idioms 71–2; idioms and cohesion 91–101; informational idioms 70–1; modalizing idioms 73–4; multiple functioning 74–83; organizational idioms 72–3
 “discourse entities” 57
 “discourse referents” 57
- embedding 140–6; adjunct clauses 144–5; complement clauses 143–4; relative clauses 141–3; relevant idioms
 distributions over embedding types 141
 emblems 53; special cases 53–5
 evaluative idioms 67, 71–2
 evaluative multifunctional idioms 76–83; evaluative-modalizing idioms 81; informational-evaluative idioms 76–81; informational-evaluative-modalizing idioms 81–3
extragrammatical idioms 38
- figurativity* 86
 fixed expressions and idioms (FEIs) 72
functional expressions 5–6
- global motivation* 38
 grammar: active and passive voice 122–7; Arabic idioms 113–49; idioms
 inflectability 113–22; idioms syntactic behaviour 127–33
 grammaticalization 54
- halub* *see* “milch”
 ‘*have a good nose*’ 104
 hyperbole 52–3
- ideational idioms* 83–4
idiom 4, 15–16, 150–5; base form 15; cohesion 154; discursive behaviour 152–4; discursive functions 152–4; inflectability 113–22; isomorphism 151–2; lexical and grammatical behaviour 154–5; nouns and adjectives
 inflectability 119–22; semantic extension patterns 150–1; semantic structure 150–2; verbs inflectability 113–19
idiom-internal relativization 146
idiomatic combinations 56

- idiomatic phrases* 56
idiomatic polysemy 53
idioms with garden-path constituents 39
 informational idioms 67, 70–1; nominal
 70; prepositional 70–1; verb-subject and
 verbal 70
 institutionalization 5
 isomorphic idioms 93, 94, 125
 isomorphism 55–61, 91; idioms 151–2;
 patterns of semantic extension in idioms
 57; syntactic types of idioms 57
- jaffa* see “to dry”
- kick the bucket* 39, 104
- LDC Arabic Corpus* 23
 lexical cohesion 91
 lexical variation: adjective variation 107;
 Arabic idioms 104–9; noun variation
 105–6; preposition variation 107; verb
 variation 105–7
 lexicogrammar: Arabic idioms 109–13;
 changes in the complexity 109–11;
 perspective-adaptation 111–13
literally non-compositional idioms
 38, 40
local grammars 113
- media Arabic 25
 metaphor 43–5, 64; and metonymy 46–51;
 see also metaphonymy
 metaphonymy 46–50; metaphor from
 metonymy 46–8; metaphor within
 metonymy 49–51; metonymy within
 metaphor 48–9
 metonymy 45, 63–4; and metaphor 46–51;
 see also metaphonymy
 “milch” 59
 modalizing idioms 67, 73–4
 Modern Standard Arabic (MSA) 3, 24;
 sources of idioms in 3–4
 motivation 37–9, 91
mufraḡah 42
 multiple functioning 74–83; evaluative
 multifunctional idioms 76–83;
 non-evaluative multifunctional
 idioms 75–6
- najm* see “star”
naran see “two fires”
nominal idioms 31–2, 70
non-decomposable idioms see idiomatic
 phrases
- non-evaluative multifunctional idioms
 75–6
 non-isomorphic idioms 55–6, 59, 93–4
 noun variation: idioms 105–6
- organizational idioms 67, 72–3; cohesion
 95–6
- passive voice 122–7
pattern reforming 146
 phraseology 1
 predicatization: idioms 147–9
 preposition variation: idioms 107
prepositional idioms 32–3
pull strings 56
- qirbah* 42
- relative clauses 141–3
 relative fixedness 5
 restricted collocations 5
rock the boat 38
rummah 54
- secondary metaphorization* 52–3
 semantics: Arabic idioms 37–65;
 conventional knowledge 51–2;
 distribution of idioms over figurative
 patterns 40; emblems 53–5; hyperbole
 52–3; interaction of metaphor and
 metonymy 46–51; isomorphism 55–61;
 isomorphism and syntactic types of
 idioms 57; isomorphism and the
 patterns of semantic extension 57;
 metaphor 43–5; metonymy 45;
 motivation 37–9; patterns of semantic
 extension 39–55
shoot one’s mouth off 48–9
shoot the breeze 38
 simple repetition 91
spill the beans 56–7
 “star” 58
strain at a gnat and swallow a camel 38
swallow 38
swallow the bitter pill 38
 syntax: adjectival idioms syntactic
 structures 132; adverbial idioms
 syntactic structures 132; Arabic idioms
 127–49; constituent order 133–40;
 embedding 140–6; idioms syntactic
 types and structures 127–33; nominal
 idioms syntactic structures 131;
 predicatization 147–9; prepositional
 idioms syntactic structures 132;

syntactic structures of verb-subject
idioms 128; verbal idioms syntactic
structures 129–30
systematic formal variation 102–3

text 66

“the bait” 58

“the honey” 58–9

“the poison” 58

“to dry” 108

to go around in circles 63–4

“to set” 58

“to swallow” 58

“two eyes” 49

“two fires” 59

verb-subject idioms 16–17, 27, 70

verb variation: idioms 105–7

verbal idioms 28–31, 70; syntactic
structures 129–30

wordplay: idioms 102–3



Routledge Paperbacks Direct

Bringing you the cream of our hardback publishing at paperback prices

This exciting new initiative makes the best of our hardback publishing available in paperback format for authors and individual customers.

Routledge Paperbacks Direct is an ever-evolving programme with new titles being added regularly.

To take a look at the titles available, visit our website.

www.routledgepaperbacksdirect.com

 **Routledge**
Taylor & Francis Group

ROUTLEDGE PAPERBACKS DIRECT



A World of Online Content!



Did you know that
Taylor & Francis
has over 20,000
books available
electronically?

What's more, they are all available for browsing and individual purchase on the Taylor & Francis eBookstore.

www.ebookstore.tandf.co.uk

eBooks for libraries

Choose from annual subscription or outright purchase and select from a range of bespoke subject packages, or tailor make your own.

www.ebooksubscriptions.com

For more information, email
online.sales@tandf.co.uk

**Free
trials
available**



Taylor & Francis 
Taylor & Francis Group

eupdates

Taylor & Francis Group

Want to stay one step ahead of your colleagues?

Sign up today to receive free up-to-date information on books, journals, conferences and other news within your chosen subject areas.

Visit
www.tandf.co.uk/eupdates
and register your email address, indicating your subject areas of interest.

You will be able to amend your details or unsubscribe at any time. We respect your privacy and will not disclose, sell or rent your email address to any outside company. If you have questions or concerns with any aspect of the eUpdates service, please email eupdates@tandf.co.uk or write to: eUpdates, Routledge, 2/4 Park Square, Milton Park, Abingdon, Oxfordshire OX14 4RN, UK.

