

# On the Distinction and Presentation of Idioms and Collocations in Bilingual English-Greek Dictionaries

**Paschalia Patsala**

School of English

Dept. of Theoretical and Applied Linguistics

Aristotle University of Thessaloniki

541 24, Thessaloniki

GREECE

ppatsala@enl.auth.gr

## Abstract

The distinction between idioms and collocations has long been recognised as being problematic. In the present paper I focus on the distinction between these multi-word constructions on the basis of linguistic criteria already proposed in the framework of various approaches. I also investigate the lexicographical choices made during the compilation of two bilingual English-Greek dictionaries with regard to idioms and collocations. These choices pertain to the selection of idioms and collocations to be included, to the selection of their translational equivalents in the target language, the microstructure of entries including collocational information, and finally to lexicographical instructions such as labels. Following an overview of these choices, I suggest that idioms and collocations are more usefully distinguished based on the theoretical and practical considerations (such as frequency, semantic opacity, internal lexical or syntactic variation, etc), and show how these considerations may be applied by presenting a revised entry for bilingual English-Greek dictionary.

## 1 Introduction

The present paper investigates the role of idioms and collocations in the compilation of dictionaries, and especially of English-Greek bilingual dictionaries. The distinction between idioms and collocations is introduced on the basis of linguistic criteria already proposed in the framework of different approaches. In order to prove whether these criteria are adequate, we have examined and further tested in corpora actual entries of two bilingual dictionaries, the *Oxford English-Greek Learner's Dictionary* (1998) and the *Collins English-Greek Dictionary* (2002).

## 2 Theoretical background

Cruse (1986) was one of the first to study idiomaticity within a formal theory of semantics, claiming that collocations are "sequences of lexical items which habitually co-occur, but which are nonetheless fully transparent in the sense that each lexical constituent is also a semantic constituent" (1986: 41). Idioms, on the other hand, are expressions "whose meaning cannot be inferred from the meaning of its parts" (1986: 37), and which are translated with semantically and not necessarily lexically equivalents in the target language. Moreover, ac-

According to Cruse (1986) the majority of idioms are homophonous with grammatically well-formed and semantically transparent expressions, with the exception of 'asyntactic idioms', such as *by and large* or *far and away* (1986: 37). Cruse also claims that we cannot insert in idioms semantically transparent lexemes, nor can we replace its parts by other semantically or syntactically equivalent lexemes.

According to another theoretical approach, proposed by Fillmore, Kay and O'Connor (1988), a syntactic construction may be idiomatic in the sense that "it may specify a semantics (and /or pragmatics) that is distinct from what might be calculated from the associated semantics" of its parts (1988: 501). The innovation of this theoretical approach is the detailed classification of idiomatic phrases on the basis of semantic and syntactic criteria in the following categories:

a) *decoding* (e.g. *give sb a wide berth*) versus *encoding* idioms that language users might understand without prior experience (for example, *the spirit is willing, but the flesh is weak*).

b) *grammatical* (such as *relax your hold*) versus *extra-grammatical idioms* which violate grammatical rules (like *by and large* or *σώνει και καλά*).

c) *lexically filled idioms* (e.g. *βάζω τελεία και παύλα*) versus *lexically open idioms* without specified lexical meaning but with a given structure (for example, *the more ... , the x-er*).

For Sinclair (1991), on the other hand, lexical meaning is determined by two parameters:

a) the *open-choice principle*, according to which "language text is the result of a very large number of complex choices" (1991: 109) the only constraint being grammaticalness, and b) the *idiom principle* which is put forward to account for the constraints not captured by the open-choice principle. He also mentions the following characteristics of the idiom principle: a) not clearly defined limits of phrases, b) internal lexical or syntactic variation (for example *set x on fire*, instead of *set fire to x*), and c) possible change of word order (1991: 110).

In *Reading Concordances* (2003), Sinclair claims that the term collocation can be defined either as a highly frequent co-occurrence of two or more words, or as a combination of words which represents a fully grammatical structure, for example a noun phrase (2003: 173). Sinclair himself prefers the first way to define collocations, and he fully adopts Firth's claim that "you shall know a word by the company it keeps" (1957: 11). Finally, he regards idioms as a category of collocations and promotes the idea of a continuum (Ooi, 1998: 57), where we have: free expressions (Open-choice principle) – variable expressions – fixed expressions (Idiom principle)

Gibbs (1994), on the other hand, has adopted a psycholinguistic approach, according to which the metaphorical basis of idioms provides the link between them and their non literal meaning. In other words, the meaning of many idiomatic phrases draws on various metonymic and metaphorical schemes used during the interpretation process (in Cameron and Low, 1999: 30). Especially useful for the treatment of multi-word constructions is the use by Gibbs of the terms *degree* and *graded condition*. In this way, the distinction between idioms and collocations can be represented as a scale. At one end we have *free-standing metaphors* (for example *deep* meaning "important"), at the other end clear cases of idioms (like *let the cat out of the bag*), and in-between cases which exhibit elements of the two extreme categories (in Cameron and Low, 1999: 126, 131).

The problem of distinguishing between idioms and collocations has been a long-time

concern not only of linguists, but also of lexicographers. Landau (2001) adopts the traditional way of defining idioms as “a group of two or more words whose collective meaning cannot be divined” by a user of a language, even if he or she knows the meaning of its parts (2001: 309). Landau examines to what extent idioms have a fixed structure or whether they are subject to changes. He mentions that the phrase to *put something on the back burner* which means “to delay considering it”, is found in various forms (2001: 316), and for this reason he introduces the term *variable idioms*. With regard to collocations, Landau describes them as “habitual structures sounding natural” (2001: 309).

Jackson, on the other hand, deals with idioms as “phrasal lexemes, typically metaphorical or figurative in meaning” (2002: 6). He does not admit of the number of co-occurring words as a defining criterion for an idiom claiming that an idiom can extend from a phrase up to a whole sentence (e.g. the difference in length between *kill two birds with one stone* and *to mess about*). Jackson attempts to define what an idiom is by referring to two criteria which he calls “characteristics”: firstly, the meaning of an idiom is something beyond the sum of the meanings of its parts, and secondly it has a relatively fixed structure. With regards to collocations, Jackson describes them as the “regular, typical or particular company that a word keeps” (2002: 99), wishing to stress the fact that it is a matter of statistical frequency.

Following an overview of these theoretical approaches, one cannot claim that there is a single approach which solves all problems concerning the distinction between these multi-word constructions. Although the criteria proposed constitute a useful guide for the lexicographer, they are just ‘clues’, not adequate for this distinction. Certainly, some of the briefly described approaches, especially the traditional approach proposed by Cruse (1986) and the collocational by Sinclair (1991), are more dominant than others. Possibly because they are based on somewhat objective and measurable criteria (namely the criteria of fixed structure, frequency of use, and semantic transparency).

Therefore, feeling that the key-notion about the distinction between idioms and collocations, is ‘degree’, I have proposed a classification of multi-word constructions on a ‘continuum’:

- a) highly frequent and semantically transparent phrases (i.e. clear cases of collocations), for example *team spirit*, *constructive criticism*, *take a decision*, etc.
- b) highly frequent co-occurrences of words (namely in-between cases of idioms and collocations), such as: *have a nice day!* *answer the door*.
- c) fixed but semantically transparent constructions (some ‘hybrid’ idioms usually based on ‘dead’ or ‘frozen’ metaphors), like *to sweeten the pill*, *quiet as a mouse*, and
- d) fixed and semantically non-transparent phrases (which are clear cases of idioms).

### 3 Data

Wishing to examine whether the above theoretical criteria have been taken into consideration in real lexicographical products, I have examined seven entries<sup>1</sup> rich in idioms and col-

<sup>1</sup> The lemmata selected were the following: *fill*, *kill*, *pace*, *positive*, *sink*, *spirit*, and *wide*.

locations of the three dictionaries already mentioned. The data has been recorded in the form of tables (an example is provided below for the verb 'sink'). Here, the distinction between idioms and collocations is based on LDOCE4 (2003).

**Idioms:**

<i>sink</i>	<i>LDOCE</i>	<i>Oxford</i>	<i>Collins</i>
one's heart ~s	X	X (μεταφ)	☐ his heart sank at the thought
one's spirits ~	X		
~ing feeling	X	X	▷ I experienced that ~ing
be sunk	X		
~ without trace	X		
~ so low	X	~ low	
~ or swim	X	X (μεταφ)	
~ your differences	X	let's ~ our differences	
~ like a stone	X	X	
~ in / into (για υπέρ & μεταφορ: αστεροειδής,		X	
~ into one's mind		X	
~ing fund		X	

**Collocations:**

<i>sink</i>	<i>LDOCE</i>	<i>Oxford</i>	<i>Collins</i>
~ to the bottom of	X		
the sun is ~ing	X	the sun was ~ing in the	
~ into (a chair)	X		▷ he sank into a chair / the
~ to one's knees	X	he sank to his knees	☐ she sank to her knees in
be ~ing fast	X	the patient is ~ing fast	
~ sth into sth (=spend a lot of money	X		▷ To sink one's teeth into
~ into insignificance		X	
~ into vice		X	
~ into a deep sleep		X	
~ into oblivion		X	
the ground ~s to the sea		X	
(prices ~)		Prices are sinking	
(the patient ~s)		The patient is sinking	
to ~ in thought/despair		X	

Figure 1. Table of idioms and collocations of the verb 'sink' in the three dictionaries.

**4 Analysis of Data**

The main problem I have faced during my analysis was the absence of consistency in lexicographical choices made by the compilers of the above English-Greek dictionaries concerning idioms and collocations.

**4.1 Oxford English-Greek Learner's Dictionary (1998)**

More precisely, with regard to the *Oxford English-Greek Learner's Dictionary*, according to its preface it is a "dictionary of modern idiomatic English language" (1998: iii). Therefore,

part of it should be the recording of idioms and collocations. Idiom in the preface is defined as “a phrase whose meaning is difficult or impossible to guess from the meanings of the individual words in it” (1998: viii). In other words, the main criterion for distinguishing between idioms and collocations here is semantic opacity.

Throughout my research I have noticed that the criterion of semantic constituents as opposed to lexical ones proposed by Cruse (1986) is consistently applied in the *Oxford English-Greek Learner's Dictionary*, providing users with semantic and not lexical equivalents in the target language, for example: *have a bee in one's bonnet* ⇒ έχω μια έμμονη ιδέα (lit. “to have a fixed idea”) (1998: 41).

However, a problem I have faced is related to the label ‘idm’ (“idiomatic”) which is not used in a systematic way throughout the dictionary, thus causing great trouble to its users. More specifically, I have noticed that clear cases of idioms are not preceded by the label (*idm*), for example:

(1) (a) *hold out the olive branch* ⇒ κρατώ κλάδον ελαίας (lit. “hold an olive branch”) (1998: 459)

(b) *of the first water* ⇒ άρίστης ποιότητας (lit. “of excellent quality”) (1998: 840)

as opposed to undisputed cases of collocations, such as those of the lemma *decision* (i.e. *make/take/come to/arrive at/reach a decision*) which are labelled *idioms*.

Another problem I have encountered has been the use of the label ‘μεταφ’ (“metaphorical”) in order to stress the metaphorical origin of some idioms, as opposed to others that should also be preceded by the same label. For example, phrases such as *be poles apart*, which is rendered into Greek as “είμαστε διαμετρικά αντίθετοι” carries the label (*μεταφ.*), as opposed to *open one's heart to sb* (translated in Greek as “ανοίγω την καρδιά μου σε κπ”). Especially interesting in the *Oxford English-Greek Learner's Dictionary* is the case of the lemma *dust* (1998: 177), where some of its idioms carry the label (*μεταφ.*), whereas others not, although they should. For example:

(2) (a) *kick up/raise a dust* ⇒ χαλώ τον κόσμο (lit. “to spoil the world”)

(b) *throw dust in sb's eyes* (*μεταφ.*) ⇒ ρίχνω στάχτη στα μάτια κάποιου (lit. “to throw ash in sb's eyes”)

#### 4.2 Collins English-Greek Dictionary (2002)

The second dictionary under examination has been the *Collins English-Greek Dictionary*, which although recently published (2002), provides its users with much less collocational information than that provided by *Oxford English-Greek Learner's Dictionary*.

As it is always the case with bilingual (printed) lexicography, due to space restrictions only a limited number of multi-word constructions can be lemmatised. However, in the case of *Collins English-Greek Dictionary* highly frequent idioms and collocations of the English language, for example *under the bridge*, *the other side of the coin*, and many others, have not been included. Equally significant is also the absence of frequent collocations (for example in *decision* we find only the collocation *make a decision*, whereas *reach/take/come to/arrive at a decision* have been omitted).

Moreover, lexicographers of the above dictionary preferred to use the terms *phrases* and *examples*, instead of *idioms* and *collocations*. Phrases show essential constructions and us-

age, they are in bold fonts and preceded by a little triangle (▶). Examples, on the other hand, (which have been derived by the Bank of English) represent the use of the specific sense in real speech, they are printed in italic bold type and are preceded by a box (☐). The problem however appears if one tries to match the pair *phrases-examples* with the pair *idioms-collocations*. Unfortunately such a matching seems to be impossible.

## 5 Concluding remarks – Suggestions

Having examined the data, I have noticed that the number of multi-word lexemes and their classification varies among dictionaries. Having in mind the theoretical criteria already proposed both by linguists and lexicographers, I have adopted the following criteria as the most helpful for the distinction between idioms and collocations in bilingual dictionaries: a) semantic transparency and high frequency of use for collocations, and b) semantic opacity and fixed structure for idioms.

As a result, I suggest that collocations in bilingual dictionaries should be lemmatised according to their frequency of use in corpora. Especially in cases of polysemous words they could be included under the appropriate sense, facilitating sense distinction. I have reached the conclusion that the use of whole sentences for presenting collocations reduces the number of collocations inserted and the linguistic information provided to users. On the contrary, presenting collocations in word combinations with their equivalents in the target language – a lexicographical choice followed in the *Oxford English-Greek Learner's Dictionary* – is a brief and user friendly practice.

Finally, with regard to idiomatic expressions, one must admit that the number of idioms which can be included in bilingual editions is also limited in comparison to monolingual ones. However, it is due to their semantic opacity that idioms must be lemmatised in order to cover both encoding and decoding needs of dictionary users. It would be especially useful idioms to be presented at the end of each entry, and to be followed by their semantic – and not lexical equivalent – in the target language. Also, idioms should be recorded with the minimum number of their elements, with all other optional parts into parenthesis.

On the basis of the above lexicographical choices, the lemma 'bird' of the *Oxford English-Greek Learner's Dictionary* (1998) has been revised in order to be included in a bilingual English-Greek learners' dictionary:

### 5.1 Revising the lemma 'bird'

#### **bird** /βɜːd/ n (C)

1. πουλί 2. (καθα.) άνθρωπος, τόπος: *He's a queer/tram* ~, είναι περίεργος τόπος, *a cunning old* ~, παμπόνηρος άνθρωπος 3. (ιδίω) *A – in the hand is worth two in the bush*, (παροιμ.) κίβλιω πέντε και στο χέρι, κυρία δέκα και καρτέρα. *kill two* ~s *with one stone*, (παροιμ.) μ' ένα σμάλτο δυο τριτόνια. *give sb/get the* ~ (χωδ.) γουχαίλο κη/γουχαίλομα. *a* ~'s eye view, πανοραμική άποψη. *birdie* n (για ή σε παιδί) πουλάκι. *goal* ~, τρόφιμος των φιδασκών.

(*Oxford English-Greek Learner's Dictionary*, 1998: 47)

#### **bird** /βɜːd/ n (C)

1. πουλί *wild/migratory/wading* ~s άγρια/ αποδημητικά/καλοφάτια πουλιά, ~s *of prey* αρπακτικά πουλιά 2. (καθα.) άνθρωπος, τόπος: *He's a wise/funny/weird* ~ είναι γνοιαπιάς/αστερός/περίεργος τόπος 3. (καθα.) κοπελιά, μενούλα 4. *do* ~ *κάνω φιδασκή*, *goal* ~, τρόφιμος των φιδασκών. 5. *kill two* ~s *with one stone* μ' ένα σμάλτο δυο τριτόνια. (*have*)

καίρουχαίζομαι. *a* ~'s eye view, πανοραμική άποψη. *birdie* *n* (για ή απ. παιδί) πουλάκι, *goal*→, τρόφιμος των φυλακών.

(*Oxford English-Greek Learner's Dictionary*, 1998: 47)

**bird** /β3:δ/ *n* (C)

1. *bird* *wild* /migratory/wading → άγρια/ αποδημητικά/καλοαέτια πούλι, → *s of prey* αρπικτικά πουλιά 2 (*καθου.*) άνθρακος, τίπος; *Ho's a wise/simny/wild* ~ είναι γλωσσικό/αστείο/παρέρτος τίπος 3 (*καθου.*) κοπελάι, μανούλι 4. *do* ~ κάνει φυλακή, *goal*→, τρόφιμος των φυλακών. 5. *kill two ~s with one stone* μ' ένα σμπάρο δυο τριγόνια. (*have*) *a ~'s eye view (of)* (έχει) πανοραμική θέα. *A ~ in the hand is worth two in the bush* κόλλω πίντε και στο χέρι, παρά δάκι και καρτέρα. *free as a ~* ελεύθερο πουλι. *~s of a feather (flock together)*, εάν δεν ταριάζαμε, δε θα συμπαιρεύαμε/όπως ομοίω αεί πλάζα. *A little ~ told me* μιν το 'πε ένα πουλάκι. → *of ill omen* άγγελος κακών (ειδήσεων). *It is (strictly) for the ~* είναι για πέταμα/για τα σκουαίδια/ άχρηστο. *The early ~ catches the worm* όποιος άρπάξτε τον Κόριο ειδή, το από ζυμνάει τον αγριότη. *give sb/get the ~* (χασ.) γιωχαίζο καίγουχαίζομαι.

Finally, following the theoretically informed analysis of this paper, we could claim that lexicography can usefully accommodate the theoretical distinction between idioms and collocations, and that the benefits of doing so outdo any problems and difficulties that this distinction raises for both lexicographers, as well as for dictionary users.

**References**

**A. Dictionaries**

*Collins English-Greek Dictionary*. (2002). Second edition. [First edition 1997]. Glasgow, Harper Collins Publishers.  
 Jackson, H. (2002), *Lexicography: An Introduction*. London, Routledge.  
 Landau, S. I. (2001), *Dictionaries: The Art and Craft of Lexicography*. Second edition. [First edition 1989]. Cambridge, Cambridge University Press.  
 Ooi, V. (1998), *Computer Corpus Lexicography*. Edinburgh, Edinburgh University Press.  
 Stavropoulos D.N. and A.S. Hornby. (eds.) (1998) *Oxford English-Greek Learner's Dictionary*. Second edition. [1st ed.: 1977]. Oxford, Oxford University Press.  
 Summers, D. (ed.) (2003), *Longman Dictionary of Contemporary English*. (Fourth Edition). Harlow, Longman. (LDOCE4)

**B. Other Literature**

BNC SARA version 0.941 (1995-7), Chancellor, Masters and Scholars of Oxford University.  
 Cameron, L., Low, G. (1999), *Researching and Applying Metaphor*, Cambridge, Cambridge University Press.  
 Cruse, D. A. (1986), *Lexical Semantics*. Cambridge, Cambridge University Press.  
 Fillmore, C.J., et al. (1988), 'Regularity and idiomaticity in grammatical constructions: the case of let alone', *Language*. Vol. 64, 3, pp. 501-511.  
 Firth, J.R. (1957), 'A synopsis of linguistic theory, 1930-1955', *Studies in Linguistic Analysis*. Special Volume, Philological Society, pp. 1-32.  
 Gibbs, R. W. (1994), *The Poetics of Mind. Figurative Thought, Language and Understanding*, Cambridge, Cambridge University Press.  
 Sinclair, J. (1991), *Corpus, Concordance, Collocation*, Oxford, Oxford University Press.  
 Sinclair, J. (2003), *Reading Concordances*, London, Pearson.