Laura CIGNONI, Istituto di Linguistica Computazionale, CNR, Pisa.
Stephen COFFEY, Centro Linguistico Interdipartimentale, Università di Pisa.

A corpus-based study of Italian idiomatic phrases: from citation forms to 'real-life' occurrences

Abstract

In the current paper we present a typological description of the alternative forms and variations which Italian idiomatic phrases were found to take in a corpus of contemporary written Italian. We were concerned above all with lexical variation, both regularly occurring and occasional, not with the syntactic flexibility of idioms. Flexible search parameters were used in order to locate as many variations as possible, including alternative lexical components, shortened forms of idioms, adaptation of underlying metaphors, and alternative syntactic realizations. We relate our findings to lexicographical description.

Keywords: corpora, idioms, Italian, lexicography, variability

1. Introduction

The purpose of the present paper is to discuss and exemplify the different forms that Italian idiomatic phrases were found to take in a corpus of contemporary written Italian, and to relate our findings to lexicographical description.

1.1. Phraseological types

The phraseological units discussed in the present paper are lexical units which were judged to show a relatively high degree of idiomaticity. The items in question were drawn from the following lexicographical sources: the phraseological collections of Lapucci, Pittàno, Quartu, and Turrini *et al.*, and one general Italian language dictionary (Zingarelli). The notion of 'idiomaticity' is by no means standardized in linguistic description. For an overview of what 'idiom' has meant to different authors, see Casadei (1995) and Fernando (1996:1–23). Two studies dealing with Italian idiomatic expressions in particular are the syntactic analysis by Vietri (1990), and the semantic study by Casadei (1996).

In our own work we assessed the degree of idiomaticity by taking into account above all a phrase's position on the following two clines: (i) semantic transparency → opaqueness of the phrase, and (ii) concreteness → abstraction of the object or concept expressed by the phrase. The higher the degree of opaqueness and abstraction, the greater was judged to be the idiomaticity. In terms of Makkai's structural analysis of English idioms, the Italian idiomatic phrases we examined are all 'lexemic idioms', and mostly fall within the class termed 'tournure idioms'. There are also a significant number of 'phrasal compound idioms' - though limited to those which are semantically opaque and more abstract in nature - together with a small number of 'irreversible binomial idioms' (see Makkai, 1972:148–168, 311–339).

1.2. The corpus consulted

The corpus consulted was the Italian Reference Corpus (hereafter IRC), located at the Institute of Computational Linguistics in Pisa. The IRC is an untagged corpus of contemporary written Italian. In the period in which we consulted the IRC, the corpus consisted of approximately 15,000,000 occurrences. Newspaper and magazine articles accounted for about two thirds of the total volume. Of the remaining 5,000,000 words, over half consisted of works of fiction, and the remainder of various works of non-fiction, including technical research reports. For further information on the composition of the IRC, the reader is referred to Bindi *et al.* (1991).

1.3. Search methodology

Since the IRC is an untagged corpus and since its internal dictionary consists only of single orthographic words, the user wishing to make searches for specific multiword units (hereafter MWUs) has only limited automatic help. This is a disadvantage if one is interested in making quick searches for standard forms of expressions, but is not of great consequence if one is concerned with retrieving as many variant forms as possible. Programs designed to look for MWUs can as yet handle only a certain amount of variation.[1]

The retrieval of idioms from the IRC involved, essentially, making searches for one or more key words for each idiom and subsequently editing out irrelevant material with a word-processor. To describe the search methodology in such simple terms, however, would be to overlook the continual decisions which had to be made about what exactly to look for. Idiomatic expressions are not single words, variable only in terms of morphology and, occasionally, spelling. Many of them have more than one basic form and all could, theoretically, be personalized in one way or another by the language user. As Gibbs *et al.* put it, ".... it may be possible to lexically change any idiom in a creative manner and still be understood if there is sufficient pragmatic context" (1989:66). Search parameters, therefore, must be as flexible as possible if one is interested in retrieving virtually all occurrences, and must be continually adjusted to the needs of each idiom.

The search program used was DBT (Database Testuale), which offered a number of features useful in the retrieval of multiword units. Firstly, it enables word families to be looked for as well as single words. Secondly, the maximum distance between the single elements comprising a word family may be set by the user. Thirdly, the program is flexible with regard to word order in a word family: one can either search for a combination of elements irrespective of which comes first, or stipulate in which order they must occur in order to be retrieved. This feature is especially useful in a language such as Italian which has relatively flexible word order. Finally, contexts of user-specified length (with a maximum of approximately 1,000 words) may be retrieved. In addition to these variables, 'traditional' one-line alphabetical concordancing is also available. For a fuller description of the program, the reader is referred to Picchi (1991). Two examples of the types of searches made illustrate the types of corpus occurrences which might have been overlooked without flexible search parameters:

(i)   Basic idiom: UN PESCE FUOR D'ACQUA [a fish out of water[2]]. Word(s) looked for: (FUOR or FUORI) & (PESCE or PESCI or ACQUA). Other search parameters: variable word order and maximum distance of 10 words between the two words looked for.

Example of occurrence found: "*Dopo Don Giulio, il prete fuor d' acqua di 'La messa è finita'.....*" [After Don Guilio, the 'priest' out of water in 'The Mass is Ended'....].

(ii)   Basic idiom: FORZARE LA MANO [to force someone's hand]. Word(s) looked for: FORZ* & MANO. Other search parameters: variable word order and maximum distance of 10 words between the two words looked for. Example of occurrence found: "*Così che, magari forzando (ma con intelligenza) la mano, Anzieu ha potuto.......*" [So, forcing - somewhat intelligently - her hand, Anzieu was able to.....]

For a more detailed analysis of the way in which the search program was used to track down as many occurrences as possible, the reader is referred to Cignoni & Coffey (1995).

1.4. Number of idioms and occurrences

Searches were carried out for a total of 324 idiomatic phrases, as defined above. 47 of these were not found in the corpus. The conclusions reached in the current paper are thus based on an examination of the contextualized occurrences of 277 idioms, though it must be stated that almost half of these occurred with a frequency of 5 or less, thereby providing very limited evidence regarding typicality and variability of the individual items.

1.5. The types of variation under consideration

The types of variation discussed in this paper relate to the lexical composition of idioms. We consider an idiom to be present in more than one form if, when one compares the various occurrences, there is seen to be a substitution of one word for another, or if one or more words have been added or taken away. We examine both the alternative forms which constitute a single idiom, and those corpus occurrences in which the language user has adapted the standard form(s) of the idiom. Discussion regarding similar formal variations in English idioms may be found in Fernando (1996:42–56) and Moon (1996).

Since idiomatic variability is such a complex phenomenon, before detailing the types of formal variation encountered, we will clear the ground slightly by mentioning some phenomena which fall outside the scope of the present study. Firstly, we are not concerned with normal morpho-syntactic changes within an idiom. Nor are we concerned with instances of discontinuity or inversion in which the idiomatic phrase is merely fitting into its linguistic context. Further exclusions from the present study are standard orthographic variations of component words of an idiom, and commonly found alternations between pairs of associated words, e.g. *tra / fra* and *insieme / assieme*.

2. Types and examples of formal variation found

2.1. Alternative forms of the 'same' word

We have stated above that spelling variants have been excluded from the present study. There are some examples of orthographic variation, however, which relate specifically to MWUs rather than to individual words. An example relates to the word *cuor*, a curtailing of the word *cuore* [heart]. In everyday language *cuor* is found only within MWUs, for example *a cuor leggero* [light-heartedly]. While looking for the latter phrase, however, we also found an occurrence of *a cuore leggero*.

Another type of word-internal variation we came across is that of the use of either capital or small case letters. An example is afforded by the alternative forms *tallone d'achille* and *tallone d'Achille* [Achilles heel].

A third type of variation we will mention in this Section is the case of two verbs of identical or very similar meaning which differ formally only by virtue of the fact that one is a prefixed form of the other. The most commonly found prefix of this type was '*ri-*', which gave rise to, for example, the alternative forms *sputare / risputare il rospo* [to get something off your chest].

2.2. Alternative lexis within an idiomatic phrase

In this Section we look at a number of different phenomena, which, however, have one thing in common: in all cases alternative words sometimes appear at a certain position within the idiomatic phrase. In the first two sub-sections we distinguish between regular alternatives and occasional variants respectively. In Sections 2.2.3 and 2.2.4 semantic factors are the initial distinguishing features.

2.2.1. Regularly occurring lexical alternatives
Sometimes one of the component parts of an idiom may be realized by more than one word. For example, in the IRC the phrase *gettare acqua sul fuoco* [literally, 'to throw water on the fire' = to calm things down], is also found a significant number of times as *buttare acqua sul fuoco*. The two verbs *buttare* and *gettare* both have the basic meaning of *to throw*, and are found as alternative possibilities in this and other phrases. Another example is afforded by the phrases *il rovescio della medaglia* [the other side of the coin] and *l'altra faccia della medaglia*, which we would class together as one and the same idiom.

2.2.2. Occasional lexical alternatives
We found it useful in the case of occasional alternatives to distinguish between three sub-categories, which differ above all from a semantic point if view. The first category is that of phrases in which the substitution of one word for another has no effect on phrasal meaning. An example is the idiom *mettere il dito sulla piaga* ['to put a finger on the wound' = to touch a sore point] of which there were 10 occurrences with the verb *mettere* and one with the verb *porre*.

A second category is that in which the substitution of one word for another resulted in a slight change in the meaning of the idiomatic phrase. For example, while looking for occurrences of the idiom *dormire / riposare sugli allori* [to rest on one's laurels] (itself an example of the 'regularly occurring lexical alternatives' described in 2.2.1), we discovered that actual usage allowed greater variation than the two cited verbs. The following verbs all combined with *sugli allori* to give the basic phrase a shift of meaning: *sonnecchiare* [to doze], *vivere* [to live], *impigrirsi* [to get lazy], *campare* [to get by], and *cullarsi* [resultant phrasal meaning: 'to be complacent']. All these instances seem clearly to be paradigmatic variations of the two other verbs.

The third category consists of more extreme cases where the language user has clearly made a conscious decision to alter the phrase very significantly in order to create a particular effect, or perhaps just to play with language. The original phrase, still instantly recognizable, is overlaid with a word or words pertaining to the subject matter being discussed. The original

idiomatic phrase does not so much change its meaning as fuse its meaning with that of the word or words which have become part of the idiom. An example of this phenomenon is afforded by the phrase *un pesce fuor d'acqua* [a fish out of water], which in one occurrence becomes *il prete fuor d'acqua* [the priest out of water]. Changes can sometimes be much more complex than this. Thus the phrase *prendere il toro per le corna* [to take the bull by the horns] undergoes various transformations to become *afferrare la Bestia per le sue corna ideologiche* [to seize the Beast by its ideological horns].

2.2.3. A large set of semantically related lexical alternatives
A type of lexical alternation which needs to be treated separately from the regular-occasional distinction we have made above, is that in which a number of semantically related words are available to fill one lexical slot within a given phrase. This may occur regularly or occasionally, with or without a change of meaning. A first example is that of a verbal phrase which is usually listed in dictionaries as *navigare in cattive acque* [literally, 'to sail in troubled waters']. Corpus reality would suggest that it is the general meaning of the word *cattive* which is essential to the phrase, not that specific word. Out of 18 occurrences found, there was only one instance of *cattive*, with a total of 12 different adjectives being used in all.

A second example involves the use of a number of words which may be subdivided into three separate semantic groups. The stable part of the phrase in question is *cordoni della borsa* [the purse strings], which was found 23 times in the IRC, in each case accompanied by a verb. 10 different verbs were involved, each having one of the following three meanings: *to tighten*, *to loosen* and *to control*.

2.2.4. Verb alternation occasioned by a shift in aspectual or other focus
A final type of lexical alternative we will mention involves changes to the aspectual, or similar, focus of the verbal component of an idiom. A common finding, for instance, was that the choice between two or more verbs determined whether an idiom was being used with an inchoative or stative meaning. An example is afforded by the idiom *essere* [stative] / *scendere* [inchoative] *sul piede di guerra* [to be / go on the warpath]. Causative verbs were also often found as possible alternative components of an idiom. Thus we find the alternatives *essere fuori dei gangheri* [stative = to be furious] and *mandare fuori dai gangheri* [causative = to infuriate].

2.3. Optional elements within an idiom

We occasionally found that a phrase was used a significant number of times both with and without a particular content word. We therefore considered the latter to be an optional lexical component of the idiom. An example is the phrase *avere (tutte) le carte in regola* [literally, 'to have (all) one's papers in order'], which was found 27 times with the word *tutte* and 33 without.

2.4. The addition of words to an idiom

Occurrences which we considered from the point of view of 'addition' were, according to our typology, necessarily occasional variations since any regular addition to an idiom would qualify as an 'optional element' as described in 2.3. An example of addition is afforded by a corpus occurrence of the phrase *gettare / buttare acqua sul fuoco* [see 2.2.1] in which the adjective *fredda* is added to *acqua* thereby slightly accentuating the meaning of the phrase.

In many cases it is less clear that we should be talking in terms of the addition of words to an idiom. Words are quite often used to make the meaning of an idiom more precise by integrating into its semantic context. The words in question attach themselves to the idiom but should perhaps be viewed as remaining part of the idiom's surroundings. This often happens when a noun phrase idiom or a noun component of a verbal idiom is linked to its context through the addition of an adjective or adjectival phrase. An example from the corpus is the phrase *"l'altra faccia della medaglia tecnologica"* [the other side of the technological coin] (for further discussion of this type of modification, see Nicolas, 1995).

2.5. The omission of one or more key words

An idiom sometimes occurs in a shortened form. The presence of a key word or words guarantees that the idiom is recognized, but one or more important words are missing. An example from the IRC is the occurrence *"altro che ago nel pagliaio"* [he's no needle in a haystack], in which there is no verb indicating the concept of 'looking for'.

2.6. The selection of key words from within an idiom

By the term 'selection' we mean cases in which key words have been individually taken from the standard citation form of an idiom and then used in a different way, without however this process resulting in the 'disappearance' of the original idiom. We found this process happening especially with metaphorical idioms in which the original metaphor was foregrounded by the language user's adaptation of the idiom. An example is afforded by the idiomatic phrase *la punta dell'iceberg* [the tip of the iceberg], found in a number of adapted forms including the following two occurrences: *"solo la punta emergente di un ben più profondo iceberg"* [only the visible tip of a much deeper iceberg], and *"come se si trattasse di due punte dello stesso iceberg"* [as if we were dealing with two tips of the same iceberg].

2.7. Alternative syntactic forms

Till now, we have been considering variations above all from a lexical point of view. Certain syntactic changes may also be involved, however. For example, a recurrently found pair of syntactic alternatives was that in which a nominal phrase - forming the core of an expression - was preceded by either the verb *avere* [to have] or the preposition *con* [with]. Thus there are various corpus occurrences of both *aver l'acqua alla gola* ['to be up to one's neck in water' = to be in dire straits] and *con l'acqua alla gola.*

Another example of syntactic variation involves a lexical item within an idiom alternating with an etymological counterpart which functions as a different part of speech, thereby changing the syntactic function of the idiom. An example is that of the idiom *scoprire l'acqua calda // la scoperta dell'acqua calda* ['to discover hot water // the discovery of hot water' = to 'discover' something which was already extremely obvious].

A final case of syntactic alternation we will mention is that of a noun phrase commonly used with a support verb, for example *doppio gioco // fare il doppio gioco* [a double-cross / to double cross]. Here we could consider both phrases as constituting the same idiom.

3. Discussion and implications for lexicography[3]

3.1. Corpus frequency

The first conclusion to be drawn from our study of idiomatic variability is that very large corpora, or perhaps corpora consisting of certain types of language, will be needed in order to obtain sufficient evidence to be able to make judgments as to what constitutes typicality for individual idiomatic phrases. In our own work, we found that a not insignificant number of idioms were absent from the corpus we were consulting. Also, there were low numbers of occurrences for many idioms (in almost 50% of those found there were only 5 occurrences or less), with the result that it was impossible for us to make judgments about whether the variations we found were regularly occurring or only occasional. The number of occurrences necessary in order to establish typicality will of course vary from item to item.

A second point to make regarding frequency is the fact that if frequency is to be considered a criterion in selecting idioms for treatment in dictionaries, then it is important to locate as many instances of an idiom as possible, both canonical forms and variations. A useful example from our own searches is that of the phrase *scheletro nell'armadio* [skeleton in the cupboard]. There was one occurrence of the phrase in that exact form, 7 more occurrences in which, with the same phrasal syntax, either *scheletro* or *armadio* was in the plural form, and a further 13 occurrences in which the two key words were extracted from the idiomatic phrase and re-used in more personalized ways by the language user, as in the following example: *"Ogni armadio (è probabile) ha uno scheletro dentro. Ma gli armadi sono milioni, tutti uguali, tutti nuovissimi.."*. [Every cupboard has probably got a skeleton inside, but there are millions of cupboards, all identical and all brand new...].

3.2. Occasional versus typical usage

Turning now to the question of which types of variation should ideally be included in lexicographical description, we believe that lexicographical description is primarily about typicality, but that typicality involves not only the use of particular lexical items within an idiom, but also the tendency for an idiom to vary in a certain way, irrespective of the precise lexical changes involved. The 'skeleton' and the 'cupboard' which refuse to stay neatly closeted together are one such example – the tendency for the idiom to which they belong to be customized by the language user needs to be commented on. Another example we will cite involves 'occasional' lexical substitution occurring frequently enough for it to be considered a regular pattern. The phrase in question is the phrase *pomo della discordia* ['apple of discord' = bone of contention], which clearly lies behind, and is present in, the following phrases all found in the corpus: *archivi della discordia, ora della discordia, nodo della discordia, musical della discordia*, and *golfo della discordia*. There is clearly a tendency for the phrase to undergo a specific type of change, to act as a phrasal template, and that tendency needs to be described and exemplified.

3.3. The extent and diversity of variation

While analyzing the occurrences of idiomatic expressions found in the IRC, we were struck by both the number of phrases which allowed some sort of variation and the number of different types of variation involved. Of the 277 idioms for which occurrences were found, 41 occurred once only and for these idioms it is impossible to make any comment on corpus

attested variation. Of the remaining 236 idioms, however, no less than 126 were found to occur in more than one form. With regard to the types of variation found, it should be pointed out that in the current paper we have sketched out the major types of variation we came across, but there were also others which, for reasons of space, it has not been possible to describe.

3.4. Areas for further investigation

There are two areas we intend to investigate further, working above all with those idioms for which we have sufficient examples. Firstly, we would like to define more rigorously the basic form of the idioms, incorporating, where it seems appropriate, the notion of semantic fields together with, or instead of, individual lexical alternatives. Thus, the already quoted phrase *cordoni della borsa* could be defined thus: VERB OF 'CONTROLLING' *or* 'LOOSENING' *or* 'TIGHTENING' + *i cordoni della borsa*. Descriptions involving controlled flexibility on the part of the linguist may be a good way to tidy up much idiomatic variability.

Secondly, we wish to see what similarities can be detected between idioms which were found to vary in a similar way or ways. Any patterns which can be found, between the idiomatic lexemes themselves or between the linguistic or pragmatic environments of their corpus contexts, would help to render our understanding of idiomatic variability less fragmented, and provide data with which to further refine the difficult task of automatic recognition of idiomatic expressions.

4. Notes

[1]   For discussion relating to the automatic retrieval of idioms, see Blackwell, Fligelstone, Ingria, Segond *et al.*, Stock, Stock *et al.*, and Van der Linden *et al.*. See also information regarding the Defi Project at the University of Liège.

[2]   English translations are provided for most Italian lexemes and citations, though sometimes translations of the former are somewhat approximate due to either lack of adequate contextualization or absence of lexical equivalence in English.

[3]   For further discussion regarding the lexicographical treatment of idioms, see Clausén, Cobuild (pp. iv–vii), Cowie, D'Elia, Moon, and Nuccorini.

5. References

Bindi, Remo, Monica Monachini & Paola Orsolini. 1991. *Italian Reference Corpus: General Information and Key for Consultation*. Istituto di Linguistica Computazionale, CNR, Pisa.

Blackwell, Susan. 1987. "Syntax versus orthography: Problems in the automatic parsing of idioms". In Roger Garside, Geoffrey Leech & Geoffrey Sampson (eds), *The Computational Analysis of English: A Corpus-based Approach*. Longman: London.

Casadei, Federica. 1995. "Per una definizione di «espressione idiomatica» e una tipologia dell'idiomatico in italiano". In *Lingua e stile*, XXX n.2.

Casadei, Federica. 1996. *Metafore ed espressioni idiomatiche: Uno studio semantico sull'italiano*, Roma: Bulzoni.

Cignoni, Laura & Stephen Coffey. 1995. *Looking for preselected multiword units in an untagged corpus of written Italian: maximizing the potential of the search program DBT.* Istituto di Linguistica Computazionale, CNR, Pisa.

Clausén, Ulla & Erika Lyly. 1994. "Criteria for Identifying and Representing Idioms in a Phraseological Dictionary." In Willy Martin *et al.* (eds), *Euralex 1994 Proceedings.* Amsterdam: Free University of Amsterdam.

The Collins COBUILD Dictionary of Idioms. 1995. London: HarperCollins Publishers.

Cowie, A.P. 1981. "The Treatment of Collocations and Idioms in Learners' Dictionaries". In *Applied Linguistics*, Vol 2, No.3.

D'Elia, Caterina. 1990. "Idioms Dictionaries: Italian and English". In *Lingvisticae Investigationes,* XIV:2.

Fernando, Chitra. 1996. *Idioms and Idiomaticity.* Oxford: Oxford University Press.

Fligelstone, Steve, Paul Rayson & Nicholas Smith. 1996. *Template analysis: bridging the gap between grammar and the lexicon.* In Jenny Thomas & Mick Short (eds), *Using Corpora for Language Research.* London & New York: Longman.

Gibbs, Raymond W., Nandini P. Nayak, John L. Bolton & Melissa E. Keppel. 1989. "Speakers' assumptions about the lexical flexibility of idioms". In *Memory & Cognition,* 17 (1).

Ingria, Robert. 1995. "Lexical Information for Parsing Systems". In Donald E. Walker, Antonio Zampolli & Nicoletta Calzolari (eds), *Automating the Lexicon.* Oxford: Clarendon Press.

Lapucci, Carlo. 1993. *Il Dizionario dei Modi di Dire della Lingua Italiana.* Milano: Garzanti.

Makkai, Adam. 1972. *Idiom Structure in English.* The Hague, Paris: Mouton.

Moon, Rosamund. 1996. "Data, Description, and Idioms in Corpus Lexicography". In Martin Gellerstam, Jerker Järborg, Sven-Göran Malmgren, Kerstib Norén, Lena Rogström & Catarina Röjder Papmehl (eds). *Euralex '96 Proceedings I-II.* Göteborg University.

Nicolas, Tim. 1995. "Semantics of Idiom Modification". In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds). *Idioms: Structural and Psychological Perspectives.* Hillsdale: Lawrence Erlbaum Associates.

Nuccorini, Stefania. 1988. "The treatment of metaphorical and idiomatic expressions in learners' dictionaries". In Mary Snell-Hornby (ed.) *ZüriLEX '86 Proceedings.* A. Francke Verlag.

Picchi, Eugenio. 1991. "DBT: A Textual Database System". In Laura Cignoni & Carol Peters (eds), *Computational Lexicology and Lexicography,* Vol VII: 177-205. Istituto di Linguistica Computazionale, CNR, Pisa.

Pittàno, Giuseppe. 1992. *Frase fatta capo ha: dizionario dei modi di dire, proverbi e locuzioni.* Bologna: Zanichelli.

Quartu, B.M. 1993. *Dizionario dei modi di dire della lingua italiana.* Milano: Rizzoli.

Segond, Frédérique, Giuseppe Valetto & Elisabeth Breidt. 1996. "Formal Description of Multi-Word Lexemes with the Finite-State Formalism IDAREX". In *COLING-96 Proceedings, Vol. 2.* Copenhagen: Center for Sprogteknologi.

Stock, Oliviero. 1989. "Parsing with Flexibility, Dynamic Strategies, and Idioms in Mind", in *Computational Linguistics,* Vol 15, No.1.

Stock, Oliviero, Jon Slack & Andrew Ortony. 1993. "Building Castles in the Air. Some Computational and Theoretical Issues in idiom Comprehension". In Cristina Cacciari & Patrizia Tabossi (eds), *Idioms: Processing, Structure, and Interpretation.* Hillsdale: Lawrence Erlbaum.

Turrini, Giovanna, Claudia Alberti, Maria Luisa Santullo & Giampiero Zanchi. 1995. *Capire l'antifona: dizionario dei modi di dire con esempi d'autore,* Bologna: Zanichelli.

Van der Linden, Erik-Jan & Wessel Kraaij. 1990. "Ambiguity resolution and the retrieval of idioms: two approaches". In Hans Karlgren (ed.), *COLING-90 (Vol. 2)*, University of Helsinki.

Vietri, Simonetta. 1990. "La sintassi delle frasi idiomatiche". In *Studi italiani di linguistica teorica ed applicata*, XIX,1.

Zingarelli, Nicola. 1988. *Il Nuovo Zingarelli* (Undicesima edizione a cura di Miro Dogliotti & Luigi Rosiello). Bologna: Zanichelli.