

UNIVERSITÀ DEGLI STUDI ROMA TRE

FACOLTÀ DI LETTERE E FILOSOFIA
DIPARTIMENTO DI LINGUISTICA
DOTTORATO DI RICERCA IN LINGUISTICA SINCRONICA, DIACRONICA E APPLICATA
XIX CICLO
A.A. 2005/2006

Parole sintagmatiche in italiano

Francesca Masini

Relatore

Prof. Raffaele Simone

Coordinatrice

Prof.ssa Lunella Mereu

Tesi di dottorato europea (SSD L-LIN/01) discussa il 21 maggio 2007

Commissione: Giorgio Banti, Sergio Scalise, Nigel Vincent

Controrelatori: Christian Lehmann, Vincenzo Lo Cascio

© Francesca Masini, 2007

This dissertation deals with multiword expressions, and in particular *phrasal words*, in a constructionist perspective. The discussion is essentially based on Italian data. However, the theoretical apparatus developed is intended as a general framework that may as well be applied to multiword phenomena in other (especially European) languages.

The dissertation is divided into two parts. The first part sets the subject matter in a historical background (Chapter 1) and describes the theoretical approach adopted (Chapter 2). The second part offers an overview of phrasal words in Italian (Chapter 3) and contains three more specific case-studies (Chapters 4-6).

PART I – BACKGROUND AND THEORETICAL APPROACH

Chapter 1 – Multiword expressions. History, models and classifications

In recent years we have witnessed an increase of attention for multiword expressions. This attention especially comes from many sub-areas of linguistics, including phraseology, lexicology, and applied linguistics in general, and comprises a huge amount of different phenomena, such as idioms, collocations, metaphors, fixed expressions, formulaic expressions and so on.

However, the study of multiword phenomena traces back to (at least) the end of the XIX century: scholars such as Paul, Bréal and Sweet noted the presence of groups of words characterised by fixity and non-compositionality. Structuralism in general also gave its contribution to the study of multiword expressions. It is especially worth mentioning Charles Bally, whose description of phraseological locutions (*locutions phraséologiques*) inspired the Soviet phraseological tradition. However, despite the identification and analysis of specific multiword phenomena (such as collocations by Firth and Palmer, lexical solidarities by Porzig and Coseriu, *synapsies* by Benveniste or *synthèmes* by Martinet, syntactic compounds by Lyons, idioms by Hockett and Makkai), the study of multiword expressions did not take off. This was probably due to the fact that, in the structuralist

period, the focus was on the individuation and study of well-defined units of analysis. Multiword phenomena, on the contrary, placed themselves in between the word level and the phrase level, and therefore could not be classified in any of the canonical levels of analysis (morphology and syntax). This idea of multiword phenomena as anomalies found solid grounds in Generative Grammar, whose modular system postulated a closed set of subsequent levels of analysis. In this framework, multiword phenomena (idioms especially) were regarded as peripheral phenomena in that they violate the principle of compositionality. In the Eighties, Charles Fillmore and colleagues founded the Construction Grammar approach, whose main goal was the rehabilitation of idioms as a crucial part of our linguistic competence by means of the notion of «construction», intended as a conventionalised association of a form and a meaning. This concept was applied to idiomatic expressions, thus allowing to encode their irregular morphosemantic features and specific semantic and pragmatic properties in the construction architecture.

In the second half of the XX century, multiword phenomena have been one of the main concerns of applied approaches, such as for instance phraseology, lexicography, language teaching and psycholinguistics. A number of classifications and definitions emerged, based on both structural and semantic criteria. Further, multiword phenomena represent a challenge for computational linguistics too: on the one hand, the identification and tagging of units of analysis is not a trivial exercise; on the other, the development of advanced IT tools for the (semi-)automatic extraction of multiword phenomena from corpora constitutes an important resource for the collection of data.

In spite of this conspicuous amount of studies and results, the theoretical status of multiword expressions remained unclear and many multiword phenomena rather unexplored. In particular, Construction Grammar, whose foundations are closely related to idioms (i.e. a particular kind of multiword expression), has devoted a large part of its efforts to spreading the notion of construction to other more general domains of the grammar, leaving the issue of multiword expressions in the background. This dissertation aims at bridging this gap and providing a comprehensive theoretical and descriptive treatment of multiword phenomena in a constructionist perspective, with special attention to phrasal words.

Chapter 2 – Constructionism. Theoretical apparatus and tools

The kind of Construction Grammar developed by Charles Fillmore and Paul Kay is not the sole constructionist approach available. Quite to the contrary, Fillmore and Kay's Construction Grammar (which we might label «classical») inspired different constructionist approaches more or less connected to the original project, such as Cognitive Construction Grammar (Adele Goldberg), Radical Construction Grammar (William Croft) and Cognitive Grammar (Ronald Langacker). Other independent models recognised their closeness to Construction Grammars. These include Construction Morphology (Geert Booij), the Tripartite Architecture of the Grammar (Ray Jackendoff), usage-based approaches (Joan Bybee), usage-based models of acquisition (Michael Tomasello) and Constructions and Categories Grammar (Raffaele Simone). All these approaches share a number of basic tenets. First, they all recognise the central role of constructions as linguistic signs endowed with different degrees of complexity and abstractness. Second, they do not postulate a principled distinction between lexicon and grammar, lexicon and syntax, system and discourse. Third, they stress the importance of functional (i.e. semantic and pragmatic) factors. Of course, there are also a number of controversial points, such as: the focus on formalisation *vs.* focus on functional, cognitive and usage-based aspects; the nature of inheritance hierarchies, which constitute the main representational tool for the relations between constructions; and, finally, the postulation of categories as primitives. In the light of this complex picture, we decided to develop an approach which would reflect our view of constructionism and which we called «mild constructionism». Mild constructionism shares the basic tenets listed above, whereas it takes a stand with respect to each of the controversial points, and more precisely: it focuses on functional, cognitive and usage-based aspects of language; it makes use of default inheritance hierarchies (following Goldberg), and postulates the existence of categories as primitives (as opposed to Croft).

The rest of Chapter 2 applies mild constructionism to the study of multiword expressions and phrasal words. First of all, it is claimed that multiword phenomena are underrepresented in the lexicon-syntax *continua* proposed in various constructionist works, in which only specified and semi-specified idioms are cited. Therefore, a more

complete picture is offered, using two criteria: complexity and specification. From the point of view of function, multiword expressions have been divided into two main groups: pragmatic-discoursal multiword expressions, which rule the interaction between speakers (formulae) and/or express pieces of shared knowledge and culture (sayings); lexico-grammatical multiword expressions, which convey lexical (phrasal words) and grammatical (periphrases) meanings. Phrasal words are therefore defined as constructional strategies that are aimed at enriching the lexical categories, or word classes, of a language. Phrasal words have been further classified in a series of sub-types according to structural and functional criteria. From the structural viewpoint, we made use of tools such as the levels of representation and the degree and kind of specification. Functionally speaking, we proposed three criteria: first, we identified phrasal words with different kinds of motivation within the system; second, we distinguished between content-phrasal words and function-phrasal words; lastly, we applied the system-discourse distinction to phrasal words, thus obtaining systemic phrasal words and discoursal phrasal words. At the end of the chapter, the relationship between phrasal words and collocations is also discussed.

PART II – CASE-STUDIES

Chapter 3 – Phrasal words in Italian. Typology and classifications

In this second part of the dissertation, the theoretical apparatus developed in the previous chapter is applied to the analysis of phrasal words in Italian. The tools and criteria identified allowed us to give a fresh look at these constructions in Italian and to propose new classifications for the different lexical categories, mainly nouns and verbs, but also adjective, adverbs and other kinds of modifiers.

The last three chapters are dedicated to the theoretical and descriptive analysis of three specific kinds of phrasal words in Italian. The three case-studies proposed also aim at bridging descriptive gaps of Italian.

Chapter 4 – Case-study I. Coordinated binomials

Chapter 4 deals with coordinated binomials, i.e. structures formed by two lexical items belonging to the same lexical category and united by a link or conjunction (e.g. *bianco e*

nero lit. white and black ‘black and white’, *anima e corpo* lit. soul and body ‘body and soul’, *punto e virgola* lit. dot and comma ‘semicolon’). The chapter contains a detailed description of the morpho-syntactic and semantic properties of Italian binomials based on a corpus of approx. 500 elements. Binomials have been analysed as constructional phrasal words that, in general, present a low degree of productivity. However, binomials also prove to display some areas of sub-regularity. The co-presence of constructions endowed with a certain degree of productivity and completely idiomatic instantiations can be accounted for by the use of inheritance hierarchies and different types of specification mechanisms. Finally, binomials have also been compared with phenomena such as syntactic tight coordination on the one hand, and compounds and reduplications on the other. Binomials have proved to place themselves in an intermediate position in-between these phenomena.

Chapter 5 – Case-study II. Phrasal verbs

Chapter 5 deals with phrasal verbs, i.e. verbal constructions formed by a verbal base and a locative particle (e.g. *venire giù* lit. come down ‘to come down, to descend’, *portare via* lit. take away ‘to take away’, *mettere sotto* lit. put under ‘to put (something) under, to run over’). Phrasal verbs are quite common in Germanic languages, but represent a peculiarity in Romance. For this reason, they have long been neglected by both scholars and lexicographers. Italian phrasal verbs have been analysed as constructional phrasal words which act as an enrichment strategy of a particular kind of word class, i.e. motion verbs. Of course, there also exist metaphorical and completely idiomatic cases, which can be accounted for by the use of different types of inheritance links. Finally, we carried out a diachronic analysis of Italian phrasal verbs. The presence of these constructions in Ancient Italian texts allowed us to identify the motivation behind their emergence: the development of phrasal verbs can be traced back to a set of typological-structural changes that took place during the passage from Latin to Italian, together with a functional urge to fill a semantic domain (i.e. spatiality in the verbal system) which was left uncovered by these changes.

Chapter 6 – Case-study III. Adnominal approximators

Chapter 6 deals with adnominal approximators, i.e. modifiers like *sorta di* ‘sort of’ and *specie di* ‘kind of’, which arose from a process of reanalysis: [[Det N1] [*di* N2]] > [Det [N1 *di*] N2]). Approximators have been analysed as semi-specified phrasal words with a lexical slot to be filled by a semi-closed set of nominal items (e.g. *specie* ‘kind’, *sorta* ‘sort’, *sottospecie* ‘subspecies/lower species’). The structural and semantic properties of this kind of modifiers have been extensively described on the basis of data extracted from spoken and written corpora of Italian. Finally, approximators have been inserted into a network of constructions that allowed us to formalise and account for their relationship with other types of constructions, such as the identifying construction (e.g. *a species of elephant*) and the affective construction (e.g. *a sweetheart of a girl*).

The overview of phrasal words in Italian and the three case-studies have shown that multiword phenomena are widespread in Italian and may serve a number of different functions. Further, the use of the notion of construction and the rehabilitation of multiword phenomena have a remarkable theoretical consequence concerning the identification of levels and units of analysis. In our view, multiword phenomena define a new structural level of analysis that places itself in-between morphology (intended as the study of the structure of words) and syntax (intended as the study of the structure of free combinations of words). Further, the multiword level, like the morphological level, correlates with different functions, since it may codify both lexical and grammatical categories, giving rise to phrasal words and periphrases respectively. Categories are therefore cross-constructural abstract conceptual entities, which interact with the different structural levels in a not biunivocal way.