

Frageografia basata sui corpora per  
apprendenti di italiano /  
*Corpus-based phraseography for  
learners of Italian*

Elmar Schafroth

Institut für Romanistik, Heinrich-Heine-Universität Düsseldorf

**APPROCCI FRASEOLOGICI ALLA LESSICOGRAFIA DI APPRENDIMENTO (PHRASALEX)**

19-20 SETTEMBRE 2019, UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO-EMILIA

# Structure

- brief outline of the project
- **methodology** and theoretical background
- **motivation** and target groups
- **structure of the phraseoframes**
- **case studies**
- challenges and perspectives

## brief outline of the project

- financed by the *Deutsche Forschungsgemeinschaft*
- University of Düsseldorf, Romanistik IV
- 2018-2021 (preparation work since 2014)
- staff: director, coordinator, 2 PhDs, 3 assistants, computer specialist, 2 lexicographers, external language expert
- online database
- aim: exhaustive description of the 600 most common verbal idioms in Italian

# methodology and theoretical background



## methodology:

- 600 most common verbal idioms of Italian
- starting point: inventory in Quartu/Rossi (2012)
- selected by corpus-based frequency analysis (and judgments of native speakers)
- strictly usage-based approach (corpora: Paisà corpus, CORIS, La Repubblica Corpus, itTenTen16, Webbit)
- formal and functional description on the basis of linguistic usage

# methodology and theoretical background



## methodology:

- learner-appropriate concept
- linguistic description mainly in prose
- linguistic information is differentiated according to learner levels: basic and detailed knowledge

## technical features:

- digital representation / alfab., onomas. search
- hyperlinks to video files to illustrate prosody and gesture
- German an Italian version

# methodology and theoretical background



## theoretical background:

- Schafroth, Elmar, (2015): Italian Phrasemes as Constructions. How to Understand and Use Them. In: *Journal of Social Sciences* 11, 3, 317-337. DOI: 10.3844/jssp.2015.317.337.
- Imperiale, Riccardo/Schafroth Elmar (2019), *Frasesologia italiana basata sull'uso: lessicografia digitale per apprendenti tra la Frame Semantics e la Grammatica delle Costruzioni*. In: *Italiano LinguaDue* 11, 1, 1-28. DOI: 10.13130/2037-3597/11872.
- Schafroth, Elmar/Imperiale, Riccardo (2019): *Gebrauchsbasierte Phraseologie des Italienischen: Digitale Lexikographie zwischen Frame-Semantik und Konstruktionsgrammatik*. In: *Lexicographica* 35, 87-121. DOI: 10.1515/lex-2019-0004.

# methodology and theoretical background



## theoretical background:

- stimulated by the following theoretical approaches:
  - Fillmore's semantics of understanding and Busse's explicative semantics: **epistemically relevant knowledge about idioms** (prototypical roles, contexts, effects, etc.)
  - Construction Grammar: description of all characteristics of form and meaning (Croft 2001): **syntactic**, morphological, phonological / **semantic**, **pragmatic**, discourse-functional **properties and restrictions**

# methodology and theoretical background



## theoretical background:

- stimulated by the following theoretical approaches:
  - Construction Grammar and Frame semantics:
    - idioms are part of a frame
    - organization of idioms by semantic fields (besides alphabetical order)
    - inventory of other lexemes and phrasemes of the same semantic field





# methodology and theoretical background



## theoretical background:

- stimulated by the following theoretical approaches:
  - Construction Grammar and Frame semantics:
    - description of linguistic data in form of attribute-value structures (*phraseoframes*): knowledge types (attributes), descriptive elements (values)
    - ‚internal‘ and ‚external syntax‘

# motivation and target groups

## – motivation:

- unsatisfactory presentation in monolingual, bilingual and phraseological lexicography
- the need for an exhaustive and detailed presentation of the relevant knowledge in order to really understand and use verbal idioms in a foreign language
- objection to the widespread opinion that only native speakers can master idioms

## motivation and target groups

- **target groups:** learners and specialists
- advanced learners of Italian (especially native speakers of German)
- teachers of Italian
- university lecturers and researchers in the field of Italian language and literature
- translators and interpreters
- Italian native speakers interested in the German equivalents of Italian idioms
- publishers dealing with Italian lexicography
- anyone interested having a good knowledge of Italian

## structure of the *phraseoframes*



### linguistic knowledge about idioms

formal (and other) variants

syntax

semantics

pragmatics

further particularities

L2

dictionaries

corpus data

# structure of the *phraseoframes*

## syntax

grammatical and lexical valency

collocations of the internal syntax

cooccurring elements of the internal syntax

characteristics of the external syntax

usual clause types

# structure of the *phraseoframes*

semantics

semantic fields

meaning(s)

thesaurus lexemes

thesaurus phrasemes

# structure of the *phraseoframes*



## pragmatics

literal meaning / interpretation

situational framework

**illocutionary functions**

register

metalinguistic use

## structure of the *phraseoframes*



### further particularities

formally similar set phrases with different (not necessarily idiomatic) meanings

stylistic particularities

word formations based on idioms

cultural specificities

other kinds of knowledge about the idiom



# structure of the *phraseoframes*



L2

usage notes

(e.g. formally similar but semantically or pragmatically different idioms)

videos

# structure of the *phraseoframes*

## dictionaries

bilingual dictionaries

monolingual dictionaries

phraseological dictionaries

# Gebrauchsbasierte Phraseologie des Italienischen

(GEPHRI): <https://gephri.phil.hhu.de/>



## – case studies:

- *fare carte false* (,to do anything for someone‘)
- *voltare le spalle* (,to let someone down‘, ,...‘)
- *darsi una mossa* (,to hurry up‘)
- *essere sul piede di guerra* (,to be at loggerheads with‘)
- *dare un taglio* (,to draw a line (under something)‘, ,to put an end to s.th.‘)
- *andare nel pallone / in bambola vs. andare im tilt* (,to be completely confused‘)

# challenges

- interactive interface for users
- semantic contextual variants (polysemy)
- assignment to semantic fields

## perspectives

- linking of FRAME (Milan project) and GEPHRI (reciprocal data export)
- modelability as linked open data?
- F.A.I.R. principles
- publication form of GEPHRI also as print version?
- similar project for other languages?
- hope for better bilingual phraseography

<https://gephri.phil.hhu.de/>

Danke für Ihre Aufmerksamkeit!