

Formalising Multilingual Verbal MultiWord Expressions to Support Neural Machine Translation



**CA21167 COST Action
UniDive - (WG2 & WG3)
16-18 March 2023 -
Paris-Saclay University**

Johanna Monti and Maria Pia di Buono
UniOr NLP Research Group
Department of Literary, Linguistic and
Comparative Studies
University of Naples L'Orientale



Università degli Studi di Napoli
"L'Orientale"

Italian Clitic Verbal VMWEs

Class of VMWE made up by a verb with at least one deficient pronoun (Cardinaletti and Starke, 1999) whose semantics is not derivable from its constituent parts

- 1) *Prendersela* → 'to be offended',
- 2) *Buttarsi giù* → 'to get depressed',
- 3) *lasciarsi andare* → 'to relax'

The class of Italian pronouns lexicalised as verbal affixes is made up of -si pronouns, traditionally referred to as "reflexive pronouns" (*mi, ti, si, ci, vi, si*), the third person feminine of clitic complement pronouns (*la, le*) and two fixed pronominal Particles: *ci* and *ne*, generally having locative or demonstrative meaning.

We take into account only clitic VMWEs which either:

- do not exist without the clitic, such as *arrabbiarsi* (to be angry), *suicidarsi*, *pentirsi*, *vergognarsi*
- have a different meaning with respect to the non-clitic verb form, such as
- are constitutive elements of more complex idiomatic expressions, such as *darsela a gambe* 'cut and run', *farla finita* 'to end it', *darci dentro* 'knuckle down', *prenderci gusto* 'get a taste for'

| Type | Source | #VMWEs |
|---------------|------------|------------|
| Procomplement | De Mauro | 172 |
| Idiom | De Mauro | 239 |
| IRV | PARSEME-IT | 252 |
| Total | | 663 |

Monti, J., di Buono, M.P., Caruso, V. (forthcoming) Ontology-based formalisation of Italian Clitic Verbal VMWEs: an Approach or Supporting Machine
Translation in Monti, J. Corpas Pastor, G., and Mitkov, R. (eds.) John Benjamin Publishing Company

Italiano ↔ Inglese

prendersela
Non prendertela
per così poco.

×

take it
Don't take it so little.

🔊 🔊



```
:prendersela v it sense-to_be_offended_v_en_sense a
vartrans:Translation ;
vartrans:source :prendersela_v_it_sense ;
vartrans:target :to_be_offended_v_en_sense .

# Examples
:prendersela v it sense ex a lexicog:UsageExample;
rdf:value "Non prendertela per così poco."@it;
rdf:value "Don't be offended for so little."@en.
```