

ANNOTATING METONYMIC RELATIONS IN A CORPUS-BASED RESOURCE FOR ITALIAN VERBS

Emma Romani & Elisabetta Ježek

University of Pavia, Department of Humanities, Theoretical and Applied Linguistics Section

emma.romani01@universitadipavia.it, jezek@unipv.it

43rd Annual Conference of the German Society for Linguistics (DGfS) - Albert Ludwig University of Freiburg im Breisgau - 23-26 February 2021

1. MOTIVATION

metonymy

- the language phenomenon for which one referent is used to denote another referent associated with it
 - *He went out to drink a glass at the pub.*
 - *glass* (metonymic/source type denoting a container) refers to its content (target type, typically denoting a beverage)

annotation of metonymies

- BA research thesis (Romani, 2020)
- corpus-based perspective and methodology
- T-PAS: a corpus-based resource for Italian verbs (Ježek et al. 2014)
- metonymy vs. complex types (e.g. [Institution] [Abstract Entity, Human Group]) (Ježek & Vieu, 2014)

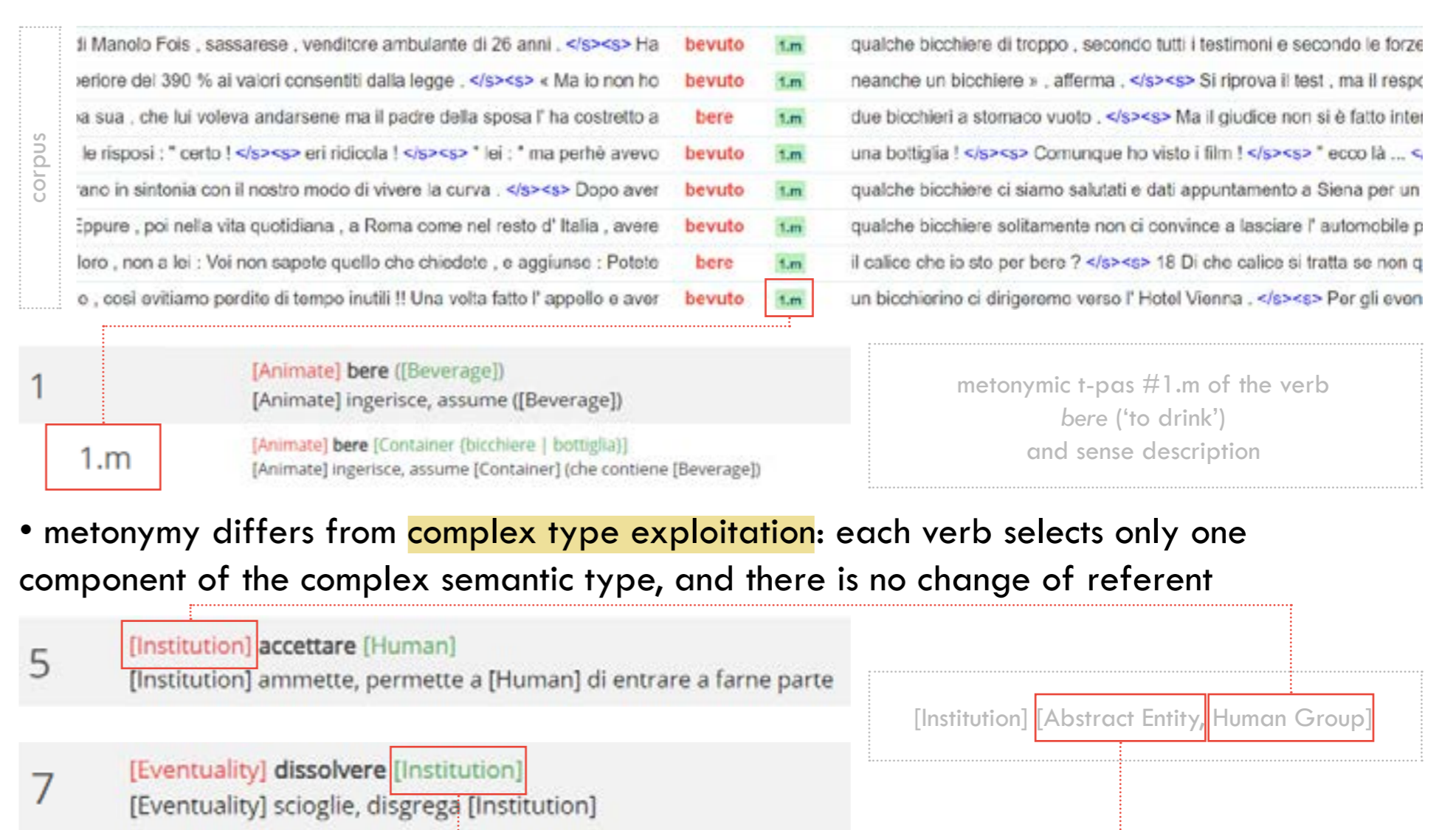
2. T-PAS: A CORPUS-BASED RESOURCE FOR ITALIAN VERBS

T-PAS

- Typed Predicate-Argument Structures
 - a repository of ca. 8,000 corpus-derived predicate argument structures for verbs with semantic specification of the arguments, linked to lexical units (called t-pas)
 - system of semantic types: an inventory of about 200 corpus-derived semantic classes for nouns, relevant for the disambiguation of the verb in context
 - a corpus of ca. 1.900,000 sentences that instantiate t-pas, tagged with lexical unit (verb) and pattern number

3. ANNOTATING METONYMIES AND COMPLEX TYPES

- 30 verbs
- dataset: 1218 annotated corpus instances
- manual annotation through Sketch Engine (Killgarriff et al., 2014) and Skema (Baisa et al., 2020)
 - random sample (200 instances)
 - instance labelling (.m)
 - sub-pattern implementation
 - definition of the metonymic relation in the sense description

4. LIST OF METONYMIC RELATIONS

relation

- a short description that illustrates how the metonymic semantic type is connected to the target semantic type
- 37 relations registered

[Machine] is activated by [Human]

The speaker announced the arrival of the train.

[Human] writes, composes [Document]

We have read Jane Austen this week at school.

[Event] is held in [Location]

I wish I could go to a concert with my friends.

[Container] contains [Beverage]

He went out to drink a glass at the pub.

[Event] happens through [Vehicle]

The flight landed at 4.30 PM in Paris.

[Human] emits [Sound]

They like to listen to Patty Smith on the radio.

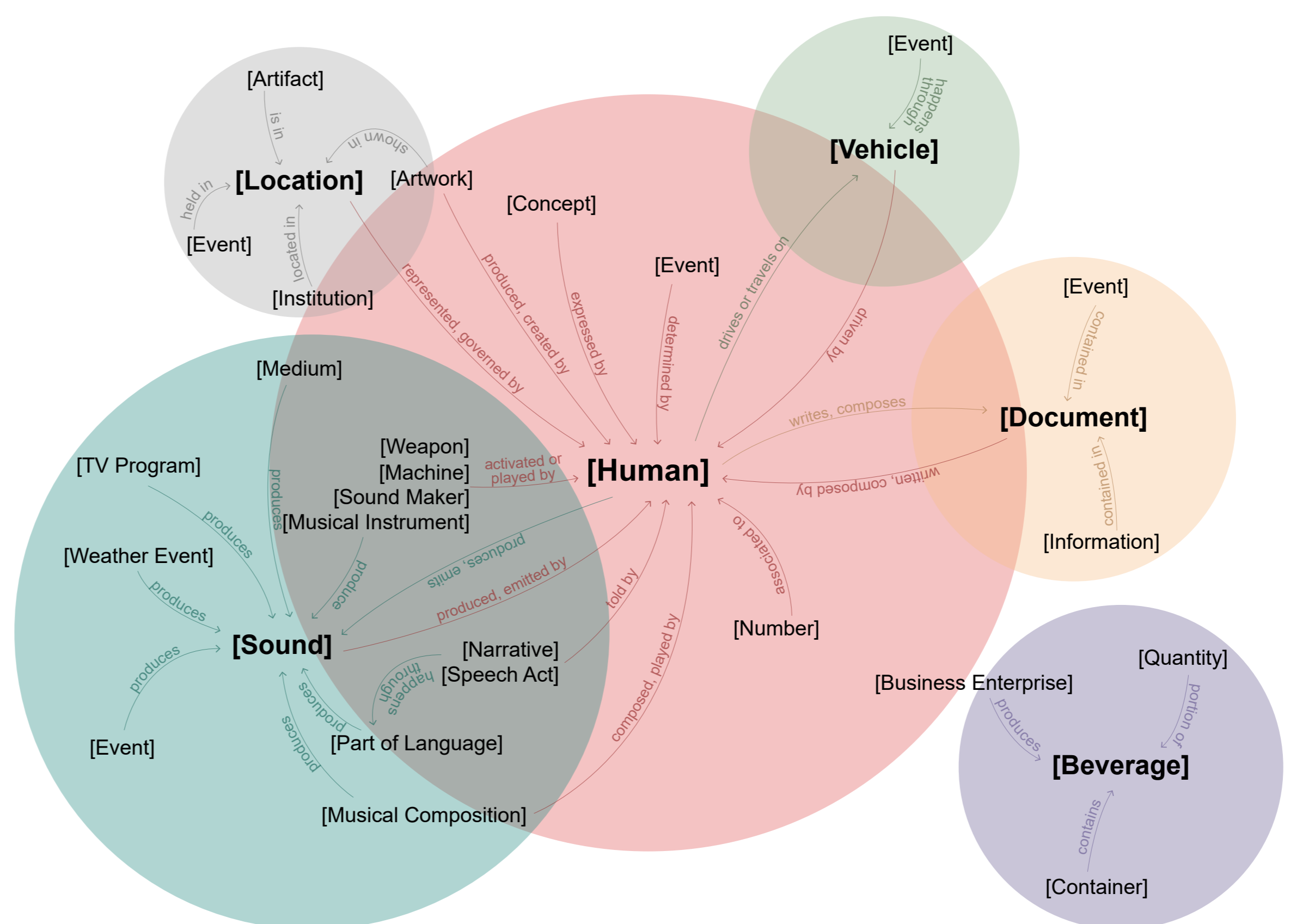
[Sound Maker] produces [Sound]

I could hear the bells ringing from my window.

[Event] of eating [Food]

The girl finished her sandwich and went away.

5. MAP OF METONYMIC RELATIONS



6. FUTURE PERSPECTIVES

- enrichment of the map and the list with new relations by extending the number of investigated verbs
- evaluation of the annotation procedure
- comparison of our results with those in Marini & Ježek (2020), in a cross-linguistic perspective
- use of our dataset and list of relations for NLP tasks, such as automatic detection of metonymies (Markert & Nissim, 2009)

REFERENCES

- Baisa, V., Tiberius, C., Ježek, E., Colman, L., Marini, C. & Romani, E. (2020). Skema: A New Tool for Corpus-driven Lexicography. In Proceedings of the 19th EURALEX International Congress.
- Ježek, E., Magnini, B., Feltracco, A., Bianchini, A. & Popescu, O. 2014. T-PAS: A resource of corpus-derived Types Predicate-Argument Structures for linguistic analysis and semantic processing. Proceedings of LREC. 890-895.
- Ježek, E. & Vieu, L. 2014. Distributional Analysis of Copredication: Towards Distinguishing Systematic Polysemy from Coercion. Proceedings of the First Italian Conference on Computational Linguistics. 219-223.
- Killgarriff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). The Sketch Engine: ten years on. Lexicography, 1(1), 7-36.
- Marini, C., & Ježek, E. 2020. Annotating Croatian Semantic Type Coercions in CROATPAS. Proceedings of the 16th Joint ACL-ISO Workshop ISA-16, pp. 49-58.
- Markert, K., and Nissim, M. 2009. Data and models for metonymy resolution. Lang Resources & Evaluation. 43: 123-138.
- Romani, E. 2020. Searching for Metonymies in Natural Language Texts. A Corpus-based Study on a Resource for Italian Verbs. BA Thesis, Pavia: University of Pavia.