Frame Semantics for Social NLP in Italian

Minnema, Gosse; Gemelli, Sara; Zanchi, Chiara; Patti, Viviana; Caselli, Tommaso; Nissim, Malvina

Published in:
Proceedings of the Eighth Italian Conference on Computational Linguistics

Frame Semantics for Social NLP in Italian: Analyzing Responsibility Framing in Femicide News Reports

Gosse Minnema[1], **Sara Gemelli**[2], **Chiara Zanchi**[2],
Viviana Patti[3], **Tommaso Caselli**[1], **Malvina Nissim**[1]
1. University of Groningen, The Netherlands
2. University of Pavia, Italy
3. University of Turin, Italy
{g.f.minnema,t.caselli,m.nissim}@rug.nl, chiara.zanchi01@unipv.it,
sara.gemelli01@universitadipavia.it, patti@di.unito.it

Abstract

We propose using a FrameNet-based approach for analyzing how socially relevant events are framed in media discourses. Taking femicides as an example, we perform a preliminary investigation on a large dataset of news reports and event data covering recent femicides in Italy. First, we revisit the EVALITA 2011 shared task on Italian frame labeling, and test a recent multilingual frame semantic parser against this benchmark. Then, we experiment with specializing this model for Italian and perform a human evaluation to test our model's real-world applicability. We show how FrameNet-based analyses can help to identify linguistic constructions that background the agentivity and responsibility of femicide perpetrators in Italian news.

1 Introduction

Frame semantics (Fillmore, 1985; Fillmore, 2006) is a theory of natural language understanding with a focus on word meanings (*lexical units*) and semantic roles (*frame elements*). The associated FrameNet project (Baker et al., 2003) has resulted in an extensive lexicon and annotated corpus implementing this theory. In the Italian computational linguistics community, there has also been considerable work on frame semantics, mostly focused on creating FrameNet resources (Tonelli and Pianta, 2008; Tonelli et al., 2009; Lenci et al., 2010; Basili et al., 2017; Brambilla et al., 2020). However the practical usability of frame semantics for Italian is still largely unexplored. First of all, on automatic *frame semantic parsing* (FSP) (Gildea and Jurafsky, 2002; Baker et al., 2007; Das et al., 2014), which has seen considerable recent work on English (Swayamdipta et al., 2017; Yang and Mitchell, 2017; Peng et al., 2018; Jiang and Riloff, 2021), there has not been any published work on Italian since the EVALITA-2011 shared task (Basili et al., 2013). Second, a clear perspective on how computational frame semantics can be useful in real-life applications is still missing.

We aim to advance the practical usability of frame semantics in Italian NLP in two ways. First, we test how well a recently developed multilingual model (LOME, Xia et al. (2021)) for FSP performs on Italian. For this purpose we use existing data from the EVALITA 2011 campaign, which is the only reference for Italian on FSP, as well as new "real world" data collected in the context of the socially relevant domain of femicides. Second, we show how frame semantics can be used in practice to run analysis on real world data. From both efforts, we draw some recommendations for practical developments in Italian FSP.

2 Semantic Frames for Events in Society

Frame semantics assumes that lexical units are points of access to complex conceptual structures: understanding the meaning of a word means to understand all of the knowledge that is associated with it. Every semantically loaded lexical item evokes a *frame*, a scenario-like unit of encyclopedic knowledge describing the concept associated to it. Frame semantics also describes the perspective in which the frame is seen. A classical example is that of a commercial transaction (Fillmore, 1971), where the same event can be presented either by foregrounding the buyer (e.g., *"Mary bought a book (from John)"*) or the seller (e.g., *"John sold a book (to Mary)"*). Perspectivization can be also related to syntactic constructions: an active sentence (*"Mary bought a book"*) and a passive one (*"The book has been bought"*) denote the same event, but make us access it via

two different participants (Meluzzi et al., 2021).

It has been shown that the variability of linguistic expressions used to describe an event impacts the reader's perception of the event and its social significance. Previous work in psycholinguistics shows that in events involving violence (at any level), the linguistic backgrounding of agents hinders their responsibility and promote victim blaming (Huttenlocher et al., 1968; Bohner, 2001; Gray and Wegner, 2009; Zhou et al., 2021; Meluzzi et al., 2021). For instance, Te Brömmelstroet (2020) shows that media in the Netherlands frequently report on traffic crashes by foregrounding the more vulnerable participants (e.g., pedestrians or cyclists), while backgrounding car drivers. A similar pattern has been observed for news reports of femicides in Italy, where the victim tends to be foregrounded and the perpetrator backgrounded (Pinelli and Zanchi, 2021; Meluzzi et al., 2021).

While there have been some proposals to use frame semantics for analyzing media framing or applying it to social media texts (Ziem et al., 2018; Brambilla et al., 2019), we are not aware of previous work that applies frame semantics to the study of linguistic perspectivization of societal issues. We test this idea and present a preliminary analysis of how frames and syntactic constructions are used to perspectivize violence in a large corpus of femicide reports in the Italian press. We adopt the *data-to-text* approach to FrameNet analysis (Vossen et al., 2020; Remijnse and Minnema, 2020; Remijnse et al., 2021), where structured event metadata is linked to texts referencing real-world events. A crucial part of this method is defining *typical frames*, i.e., frames that are hypothesized to conceptualize important aspects of the targeted event type. For the femicide domain, we selected 15 typical frames;[1] some examples are in Table 1.

3 Frame Semantic Parsing for Italian

The shared task on Frame Labeling over Italian Texts (FLAIT) at EVALITA 2011 (Basili et al., 2013) introduce the only existing published Italian FSP models, as well as the only publicly available corpus for the task on generic texts. As shown in Table 2, the FLAIT corpus contains 1,569 annotated sentences, all of which are so-called *exemplars* containing a single annotated predicate and frame structure. Compared to the English Berkeley FrameNet (BFN), which contains also fully annotated documents, the models presented at FLAIT are impressive (scores up to 80%).

3.1 LOME experiments

LOME (Xia et al., 2021) is a recent end-to-end FSP model that reports excellent frame detection scores on English, and, thanks to its XLM-R encoder (Conneau et al., 2020), is the first cross-lingual FSP model, even though it was trained on English data only. Here, we propose several strategies for adapting LOME to Italian and making maximum use of the available data.

Strategies The simplest strategy, ***LOME-EN***, is to use the English-trained model in a zero-shot setup to make predictions for Italian texts. A downside of this approach is that the model is not able to tag the Italian-specific frames that have been created in the IFrameNet project (Basili et al., 2017), which also makes the evaluation on FLAIT data more challenging. FLAIT contains 10 frames that do not currently exist in BFN (7.4% of training instances and 6.0% of test instances). It therefore makes sense to also train LOME on FLAIT directly. In ***IT-Simple***, we only train on FLAIT data; in ***IT-Concat***, we train on the concatenation of FLAIT and the fully annotated documents from BFN; and in ***IT-Berkeley***, we train only on FLAIT but initialize the encoder with the parameters of LOME-EN.

Evaluation For use in real-life applications, what truly matters is *end-to-end* performance, i.e. from raw texts to the predictions of all predicate frames and associated roles. Full end-to-end evaluation is impossible in FLAIT since only one predicate per sentence is annotated. However, we can approximate it by obtaining the full predictions from the models and then evaluate only on FLAIT gold predicates. In this way, models are penalized for missing predicates that should have been annotated (but not for overgeneration). We use the *SeqLabel* metric (Minnema and Nissim, 2021) for scoring frame and role label predictions on a token-by-token basis.

Additionally, to test LOME against the 2011 models, we reimplement the FLAIT evaluation metrics, in which models are asked to predict (i) frames given a predicate (*Frame Detection* [FD]), (ii) semantic role spans given a frame (*Boundary*

[1] ABUSING, ATTACK, CAUSATION, CAUSE_HARM, CAUSE_MOTION, DEAD_OR_ALIVE, DEATH, EMOTION_DIRECTED, EVENT, EXPERIENCE_BODILY_HARM, HIT_TARGET, KILLING, QUARRELING, RAPE, USE_FIREARM.

Frame	Description	Example
KILLING	an agent *(Killer)* actively causes the death of a patient *(Victim)*	[The man] **killed** [his wife]
DEATH	someone *(Protagonist)* dies	[The woman] **died**
DEAD_OR_ALIVE	state of someone *(Protagonist)* being dead or alive	[She] was found **dead**
CAUSE_HARM	an agent *(Agent)* actively causes a patient *(Victim)* to be hurt	[He] **stabbed** [his girlfriend]
EVENT	an unspecified event *(Event)* happens	[The dramatic events] **happened** last week

Table 1: Examples of FrameNet frames relevant for describing femicides. Semantic role names indicated in *italics*, lexical units indicated in **bold**.

		English	Italian
fulltext	sentences	5,093	0
	frame instances	29,359	0
exemplar	sentences	163,801	1,569
	frame instances	169,473	1,569
total	sentences	168,894	1,569
	frame instances	198,832	1,569

Table 2: Sentences and annotations in the English and Italian datasets.

		frames			roles		
		P	R	F	P	R	F
EN	**LOME-EN**	0.89	0.70	0.78	0.69	0.59	0.64
IT	**LOME-EN**	0.63	0.52	0.57	0.63	0.50	0.56
	IT-Simple	-0.14	0.14	-0.01	-0.14	0.16	0.00
	IT-Concat	0.21	0.14	0.17	0.10	0.08	0.09
	IT-Berkeley	-0.07	0.17	0.05	0.04	0.12	0.09

Table 3: SeqLabel scores for gold predicates. Blue: baseline, green/red: performance deltas

Detection [BD]), or (iii) semantic role labels given a frame and the role spans (*Argument Classification* [AC]).[2]

Implementation We kept LOME model and training settings the same as described by Xia et al. (2021). During testing, we noticed that 56 instances in the FLAIT test set had misspelled frame labels,[3] causing a large drop in scores. We fixed these labels, but since we do not know if the original evaluation script also did this, we report the uncorrected scores in our GitHub repository.

Results Sequence labeling performance is reported in Table 3. The zero-shot LOME-EN model achieves an F1 score of 0.57 for frames and 0.56 for roles, substantially less than IT-Concat, which gets close to scores on English (0.74 F1 on frames, 0.63 on roles). The other two Italian models have mixed results, with improvements on recall but not on precision. However, IT-Berkeley outperforms both LOME-EN and IT-Simple, showing that reusing encoder weights helps boost performance.

Turning to EVALITA-style evaluation, in Table 4[4] we compare LOME against the best system from 2011, which is based on a SVM with a tree kernel (Croce et al., 2013). The most striking result is that, on frame prediction, the 2011 winner is still king, with the LOME-EN and IT-Concat models falling short by 0.24 and 0.04 points, respectively. For semantic role prediction, results are mixed: LOME-EN has a modest but consistent improvement on both span (BD) and label (AC) prediction, while IT-Concat improves on some setups but not on others.

3.2 Evaluating Real-World Performance

We explore how robust are our models when deployed on other data. We focus on frame prediction only, a task know to be harder to adapt across domains (Hartmann et al., 2017)

Femicide annotation We deployed the LOME-EN and IT-Concat on a set of femicide news reports (see §4) with typical frames (see §2) in an end-to-end setup (i.e., without predicates as input). Out of 4,444 frame predictions, the two models disagreed in 58% of cases. Next, for a subset of 150 conflicts, we manually annotated[5] which of the two predictions is better. Table 6 shows that LOME-EN performs much better than IT-Concat, especially on two of the most frequent typical frames (KILLING and EMOTION_DIRECTED). This is largely due to predicate detection: 47% of cases where LOME-EN is better than IT-Concat are due to IT-Concat not detecting the predicate; in conflicts for predicates that both models detected, IT-Concat slightly out-

[2] As we were unable to access the original evaluation script, we have attempted to reproduce it as faithfully as possible from the description in Basili et al. (2013).

[3] In these frame names, dashes were used in place of underscores, e.g. CAUSE-HARM instead of CAUSE_HARM.

[4] We only report strict scores for BD and AC. Full tables with token-based scores are in our GitHub repository.

[5] Annotation was done by a single annotator, who is also one of the co-authors of this paper. Annotation was blind and randomized, i.e., the annotator had no way to guess which prediction came from which model.

	run 1			run 2			run 3		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
FD									
2011-best	0.81	0.81	0.81	-	-	-	-	-	-
LOME-EN	-0.24	-0.24	-0.24	-	-	-	-	-	-
IT-Concat	-0.04	-0.04	-0.04	-	-	-	-	-	-
BD (strict)									
2011-best	0.67	0.73	0.69	0.67	0.73	0.69	-	-	-
LOME-EN	0.10	0.05	0.08	0.02	0.07	0.05	-	-	-
IT-Concat	-0.09	-0.06	-0.08	-0.10	-0.06	-0.08	-	-	-
AC (strict)									
2011-best	0.48	0.53	0.50	0.51	0.56	0.53	0.70	0.70	0.70
LOME-EN	-0.01	0.02	0.01	0.09	0.13	0.11	0.16	0.16	0.16
IT-Concat	-0.02	0.00	-0.01	-0.03	0.01	-0.01	0.14	0.14	0.14

Table 4: EVALITA-2011-style evaluation. As in the original task, run 1, 2, and 3 refer to predictions with, resp., no gold inputs, gold frame inputs, and gold frame and role span inputs.

	frames			
	all	**IFN**	**BFN**	**fcd**
FLAIT/dev				
num_examples	*123*	*123*	*113*	*14*
Simple SVM	0.59	0.59	0.60	0.71
LOME-EN	0.59	0.59	0.65	0.71
IT-Concat	0.85	0.85	0.87	0.93
FLAIT/test				
num_examples	*318*	*318*	*299*	*43*
Simple SVM	0.29	0.29	0.30	0.40
LOME-EN	0.57	0.57	0.60	0.60
IT-Concat	0.77	0.77	0.76	0.81
femicides				
num_examples	*43*	*43*	*43*	*43*
Simple SVM	0.14	0.14	0.14	0.14
LOME-EN	0.63	0.63	0.63	0.63
IT-Concat	0.72	0.72	0.72	0.72

Table 5: Generalizability scores

	best prediction			
	EN	**IT**	**both**	**none**
overall	0.51	0.12	0.12	0.25
non-null	0.17	0.22	0.44	0.17
by frame				
KILLING	0.70	0.19	0.11	0.00
EMOTION_D.	0.77	0.05	0.05	0.14
DEATH	0.33	0.05	0.19	0.42

Table 6: Conflict analysis on the femicides dataset. 'EN': LOME-EN; 'IT': IT-Concat; 'both'/'none': both models are equally correct/wrong.

performs LOME-EN. We speculate that this might be explained by the exemplar-style structure of the FLAIT corpus.

Generalization Table 5 shows frame detection scores on three evaluation sets: the FLAIT development set (10% held-out from the training set), the FLAIT test set, and the set of cases from our femicide annotation experiment in which at least one of the two models' predictions was marked as correct.[6] Since we do not have access to the original FLAIT models, we use a simple linear SVM,[7] trained on FLAIT, as an alternative baseline. The task is the same as the FLAIT FD task: the models are given the gold predicate and asked to predict the frame. Results are split by frame category: IFrameNet frames that FLAIT-trained models can be expected to know ('IFN'), BFN frames that LOME-EN can be expected to know ('BFN'), and typical frames for femicides ('fcd').

The results show several patterns that are relevant for real-world usability. First, both LOME models perform as good or better on typical femicide frames compared to other frames, which is a positive sign for the feasibility of our project. Furthermore, IT-Concat is clearly the overall best frame detection model, but only when it already knows which predicates to annotate (see above). However, it is also quite biased towards the FLAIT dataset, scoring substantially worse on the test and femicide datasets compared to the development set. By contrast, LOME-EN is very stable across datasets. The SVM baseline performs surprisingly well on the development set, but much worse on the test set and extremely poorly on the femicides dataset. We interpret this as a sign of the limited coverage of the FLAIT dataset, showing that good performance on the shared task is not necessarily indicative of real-world performance.

4 Frame-Based Analysis of Femicide News

In this section, we provide a concise overview of our initial work on applying frame semantic parsing to investigate news coverage of femicides.

Dataset We perform our analysis on a private dataset collected by the CRITS research team at RAI (Radiotelevisione Italiana) and made available for use in our project. The dataset contains 2,734 news articles from 31 different Italian news sources, reporting on 937 femicides perpetrated between 2015 and 2017, along with structured in-

[6] If the annotator indicated that both predictions for a particular predicate were equally good, we randomly selected one of the predictions as the 'gold' label.

[7] The SVM takes as input a bag-of-bigrams extracted from a context window of 5 tokens before and after the predicate.

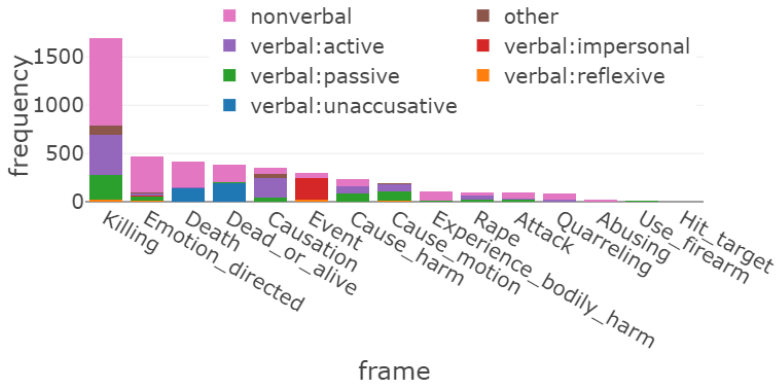

Figure 1: Typical frame frequencies, split by syntactic construction

formation about these femicides (Belluati, 2021)[8]. The dataset is unique because it includes rich event metadata, and contains various news article per femicide, allowing for investigating variation in framing of the same event along different dimensions, e.g., over time or by news source.

Analysis Based on our findings in §3, especially from the human evaluation experiment, we deploy the LOME-EN model to automatically annotate a randomly chosen 200K word subcorpus covering 10% of all events.The frame semantic annotations are enriched with dependency parses produced by `spaCy` (Honnibal et al., 2020), which are converted into syntactic construction annotations using a set of heuristics.

Figure 1 shows our main results. KILLING is by far the most frequent typical frame, followed by EMOTION_DIRECTED and DEATH. Looking at syntax, we find that *nonverbal* constructions, in which the predicate is expressed by a noun or adjective (e.g., *"l'omicidio"* "the murder") are dominant in many frames. Instead, *verbal:active* constructions (e.g., *"X uccide Y"* "X kills Y") are much rarer, as are *verbal:passive* (e.g., *"X è uccisa"* "X is killed") and *verbal:unaccusative* (e.g., *"X è deceduta"* "X has died").

Looking at semantic roles, patterns that vary greatly depending on frames and constructions. In general, semantic roles that are likely to refer to the perpetrator appear to be expressed much less frequently than those referring to the victim. For KILLING, 60% of all instances express a Victim role, while only 33% express a Killer role. However, instances with a nonverbal construction only express these roles in 40% and 20% of cases, respectively, against 71% and 87% in active constructions. On the other hand, DEATH expresses a victim-like role (Protagonist) in 79% of cases, whereas its only role that can encode a perpetrator (Explanation) occurs in 14% of cases.

While our analysis is too preliminary to draw strong conclusions, our findings are consistent with previous work: agentivity-backgrounding constructions (especially nonverbal) are very common, and semantic roles encoding the victim are more frequent than those encoding the perpetrator. What our frame analysis adds to previous work is information about the semantics of the analyzed constructions. For example, the dominance of KILLING suggests that femicides tend to be framed as agentive at least on a lexical level, even if the perpetrator is often backgrounded syntactically. On the other hand, non-agentive ways of framing the event (DEATH, DEAD_OR_ALIVE, EVENT) are also relatively common, accounting for 24% of frame instances.

5 Conclusions

We took initial steps towards addressing (i) the lack of recent frame semantic parsing models, and (ii) a missing perspective on how frame semantic analysis can be applied in practice. We adapted the multilingual LOME parser (Xia et al., 2021) to Italian, tested it against the EVALITA-2011 benchmark, and performed experiments to evaluate its real-world performance. Furthermore, we hypothesize that frame semantics can be a valu-

[8]The dataset has been collected as an outcome of the PRIN 2015 research project *Rappresentazioni sociali della violenza sulle donne: il caso del femminicidio in Italia*.

able analysis tool for analyzing backgrounding (and indirectly, blame attribution) of event participants, and propose news reports about femicides as an example of a domain where this type of analysis is very socially relevant.

Our results indicate that LOME-based models can achieve acceptable performance, both on the EVALITA benchmark and out-of-domain on femicide reports, even without a large quantity of training data. We also found that a cross-lingual approach is useful: training on the concatenation of English and Italian data yields substantial improvements over using only Italian data, and even a zero-shot approach with only English data works quite well. However, our real-world performance analysis highlights key limitations of the Italian data: while models trained on EVALITA can achieve good frame detection performance, they fail when used 'end-to-end', with predicate identification seemingly the main bottleneck.

Finally, we performed a preliminary framing analysis of a large dataset covering femicides in Italy. While our analysis method is still in very early stages, we believe that our initial results demonstrate that frame semantics is meaningful for analyzing femicides and other social issues, and that it complements earlier construction-based approaches. In the future, we aim to expand our analysis system to make it usable for different social applications: for example, one could envision systems that can help social scientists test specific hypotheses about media reporting, help activists identify and highlight biased forms of reporting, or help make journalists more aware of their writing and its possible social-cognitive effects.

Acknowledgements

We would like to thank the CRITS department at RAI for giving us access to the femicides dataset. We would also like to thank our collaborators in the broader responsibility framing research effort that this work is part of: Marion Bartl, Gaetana Ruggiero, Marco te Brömmelstroet, and Eva Kwakman. Authors G.M., T.C., and M.N. worked on this paper as part of the project *Framing situations in the Dutch language* (code: VC.GW17.083/6215), funded by the Dutch National Science Organization (NWO).

References

Collin F. Baker, Charles J. Fillmore, and Beau Cronin. 2003. The structure of the FrameNet database. *International Journal of Lexicography*, 16(3):281–296.

Collin Baker, Michael Ellsworth, and Katrin Erk. 2007. SemEval-2007 task 19: Frame semantic structure extraction. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 99–104, Prague, Czech Republic, June. Association for Computational Linguistics.

Roberto Basili, Diego De Cao, Alessandro Lenci, Alessandro Moschitti, and Giulia Venturi. 2013. Evalita 2011: The frame labeling over Italian texts task. In Bernardo Magnini, Francesco Cutugno, Mauro Falcone, and Emanuele Pianta, editors, *Evaluation of Natural Language and Speech Tools for Italian*, pages 195–204, Berlin, Heidelberg. Springer.

Roberto Basili, Silvia Brambilla, Danilo Croce, and Fabio Tamburini. 2017. Developing a large scale FrameNet for Italian: the IFrameNet experience. In Roberto Basili, Malvina Nissim, and Giorgio Satta, editors, *Proceedings of the Fourth Italian Conference on Computational Linguistics (CLiC-it 2017), Rome, Italy, December 11-13, 2017*, volume 2006 of *CEUR Workshop Proceedings*. CEUR-WS.org.

M. Belluati. 2021. *Femminicidio. Una lettura tra realtà e interpretazione*. Biblioteca di testi e studi. Carocci.

Gerd Bohner. 2001. Writing about rape: Use of the passive voice and other distancing text features as an expression of perceived responsibility of the victim. *British Journal of Social Psychology*, 40(4):515–529.

Silvia Brambilla, Alessio Palmero Aprosio, and Stefano Menini. 2019. BullyFrame: Cyberbullying meets FrameNet. In Raffaella Bernardi, Roberto Navigli, and Giovanni Semeraro, editors, *Proceedings of the Sixth Italian Conference on Computational Linguistics, Bari, Italy, November 13-15, 2019*, volume 2481 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Silvia Brambilla, Danilo Croce, Fabio Tamburini, and Roberto Basili. 2020. Automatic induction of FrameNet lexical units in Italian. In Johanna Monti, Felice Dell'Orletta, and Fabio Tamburini, editors, *Proceedings of the Seventh Italian Conference on Computational Linguistics, CLiC-it 2020, Bologna, Italy, March 1-3, 2021*, volume 2769 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In

Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 8440–8451, Online, July. Association for Computational Linguistics.

Danilo Croce, Emanuele Bastianelli, and Giuseppe Castellucci. 2013. Structured kernel-based learning for the frame labeling over Italian texts. In *Evaluation of Natural Language and Speech Tools for Italian*, pages 195–204, Berlin, Heidelberg. Springer Berlin Heidelberg.

Dipanjan Das, Desai Chen, André F. T. Martins, Nathan Schneider, and Noah A. Smith. 2014. Frame-semantic parsing. *Computational Linguistics*, 40(1):9–56, March.

Charles J. Fillmore. 1971. Subjects, speakers, and roles. *Synthese*, 21(3-4).

Charles J. Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di semantica*, 6(2):222–254.

Charles J. Fillmore. 2006. Frame semantics. In D. Geeraerts, editor, *Cognitive Linguistics: Basic Readings*, pages 373–400. De Gruyter Mouton, Berlin, Boston. Originally published in 1982.

Daniel Gildea and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288.

Kurt Gray and Daniel M. Wegner. 2009. Moral typecasting: divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, 96:505–520.

Silvana Hartmann, Ilia Kuznetsov, Teresa Martin, and Iryna Gurevych. 2017. Out-of-domain FrameNet semantic role labeling. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 471–482, Valencia, Spain, April. Association for Computational Linguistics.

Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. 2020. spaCy: Industrial-strength Natural Language Processing in Python.

Janellen Huttenlocher, Karen Eisenberg, and Susan Strauss. 1968. Comprehension: Relation between perceived actor and logical subject. *Journal of Verbal Learning and Verbal Behavior*, 7:527–530.

Tianyu Jiang and Ellen Riloff. 2021. Exploiting definitions for frame identification. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2429–2434, Online, April. Association for Computational Linguistics.

Alessandro Lenci, Martina Johnson, and Gabriella Lapesa. 2010. Building an Italian FrameNet through semi-automatic corpus analysis. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May. European Language Resources Association (ELRA).

Chiara Meluzzi, Erica Pinelli, Elena Valvason, and Chiara Zanchi. 2021. Responsibility attribution in gender-based domestic violence: A study bridging corpus-assisted discourse analysis and readers' perception. *Journal of Pragmatics*, 185:73–92.

Gosse Minnema and Malvina Nissim. 2021. Breeding Fillmore's chickens and hatching the eggs: Recombining frames and roles in frame-semantic parsing. In *Proceedings of the 14th International Conference on Computational Semantics*. `https://iwcs2021.github.io/proceedings/iwcs/pdf/2021.iwcs-1.15.pdf`.

Hao Peng, Sam Thomson, Swabha Swayamdipta, and Noah A. Smith. 2018. Learning joint semantic parsers from disjoint data. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1492–1502, New Orleans, Louisiana, June. Association for Computational Linguistics.

Erica Pinelli and Chiara Zanchi. 2021. Gender-based violence in italian local newspapers: How argument structure constructions can diminish a perpetrator's responsibility. *Discourse Processes between Reason and Emotion: A Post-disciplinary Perspective*, page 117.

Levi Remijnse and Gosse Minnema. 2020. Towards reference-aware FrameNet annotation. In *Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet*, pages 13–22, Marseille, France, May. European Language Resources Association.

Levi Remijnse, Marten Postma, and Piek Vossen. 2021. Variation in framing as a function of temporal reporting distance. In *Proceedings of the 14th International Conference on Computational Semantics*. `https://iwcs2021.github.io/proceedings/iwcs/pdf/2021.iwcs-1.22.pdf`.

Swabha Swayamdipta, Sam Thomson, Chris Dyer, and Noah A. Smith. 2017. Frame-semantic parsing with softmax-margin segmental rnns and a syntactic scaffold. *CoRR*, abs/1706.09528.

Marco Te Brömmelstroet. 2020. Framing systemic traffic violence: Media coverage of dutch traffic crashes. *Transportation research interdisciplinary perspectives*, 5.

Sara Tonelli and Emanuele Pianta. 2008. Frame information transfer from English to Italian. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, May. European Language Resources Association (ELRA).

Sara Tonelli, Daniele Pighin, Claudio Giuliano, and Emanuele Pianta. 2009. Semiautomatic Development of FrameNet for Italian. In *Proceedings of the FrameNet Workshop and Masterclass, Co-located with the Seventh International Workshop on Treebanks and Linguistic Theories*.

Piek Vossen, Filip Ilievski, Marten Postma, Antske Fokkens, Gosse Minnema, and Levi Remijnse. 2020. Large-scale cross-lingual language resources for referencing and framing. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 3162–3171, Marseille, France, May. European Language Resources Association.

Patrick Xia, Guanghui Qin, Siddharth Vashishtha, Yunmo Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron Steven White, and Benjamin Van Durme. 2021. LOME: Large ontology multilingual extraction. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pages 149–159, Online, April. Association for Computational Linguistics.

Bishan Yang and Tom Mitchell. 2017. A joint sequential and relational model for frame-semantic parsing. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1247–1256, Copenhagen, Denmark, September. Association for Computational Linguistics.

Karen Zhou, Ana Smith, and Lillian Lee. 2021. Assessing cognitive linguistic influences in the assignment of blame. In *Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media*, pages 61–69, Online, June. Association for Computational Linguistics.

Alexander Ziem, Christian Pentzold, and Claudia Fraas. 2018. Medien-Frames als semantische Frames: Aspekte ihrer methodischen und analytischen Verschränkung am Beispiel der 'Snowdon-Affäre' [Media frames as semantic frames: Aspects of their methodological and analytical entanglement in the example of the Snowdon affair]. In Detmer Wulf Alexander Ziem, Lars Inderelst, editor, *Frames interdisziplinär: Modelle, Anwendungsfelder, Methoden*, pages 155–184, Düsseldorf. DUP.