

# L'ANALISI LINGUISTICA E LETTERARIA

FACOLTÀ DI SCIENZE LINGUISTICHE E LETTERATURE STRANIERE  
UNIVERSITÀ CATTOLICA DEL SACRO CUORE

3

ANNO XXVIII 2020

EDUCATT - UNIVERSITÀ CATTOLICA DEL SACRO CUORE

L'ANALISI  
LINGUISTICA E LETTERARIA

---

FACOLTÀ DI SCIENZE LINGUISTICHE  
E LETTERATURE STRANIERE

UNIVERSITÀ CATTOLICA DEL SACRO CUORE

3

ANNO XXVIII 2020

PUBBLICAZIONE QUADRIMESTRALE

L'ANALISI LINGUISTICA E LETTERARIA  
Facoltà di Scienze Linguistiche e Letterature straniere  
Università Cattolica del Sacro Cuore  
Anno XXVIII - 3/2020  
ISSN 1122-1917  
ISBN 978-88-9335-766-1

---

*Comitato Editoriale*

GIOVANNI GOBBER, Direttore  
MARIA LUISA MAGGIONI, Direttore  
LUCIA MOR, Direttore  
MARISA VERNA, Direttore  
SARAH BIGI  
ELISA BOLCHI  
GIULIA GRATA  
CHIARA PICCININI  
MARIA PAOLA TENCHINI

*Esperti internazionali*

THOMAS AUSTENFELD, Université de Fribourg  
MICHAEL D. AESCHLIMAN, Boston University, MA, USA  
ELENA AGAZZI, Università degli Studi di Bergamo  
STEFANO ARDUINI, Università degli Studi di Urbino  
GYÖRGY DOMOKOS, Pázmány Péter Katolikus Egyetem  
HANS DRUMBL, Libera Università di Bolzano  
JACQUES DÜRRENMATT, Sorbonne Université  
FRANÇOISE GAILLARD, Université de Paris VII  
ARTUR GAŁKOWSKI, Uniwersytet Łódzki  
LORETTA INNOCENTI, Università Ca' Foscari di Venezia  
VINCENZO ORIOLES, Università degli Studi di Udine  
GILLES PHILIPPE, Université de Lausanne  
PETER PLATT, Barnard College, Columbia University, NY, USA  
ANDREA ROCCI, Università della Svizzera italiana  
EDDO RIGOTTI, Università degli Studi di Perugia  
NIKOLA ROSSBACH, Universität Kassel  
MICHAEL ROSSINGTON, Newcastle University, UK  
GIUSEPPE SERTOLI, Università degli Studi di Genova  
WILLIAM SHARPE, Barnard College, Columbia University, NY, USA  
THOMAS TRAVISANO, Hartwick College, NY, USA  
ANNA TORTI, Università degli Studi di Perugia  
GISÈLE VANHESE, Università della Calabria

*I contributi di questa pubblicazione sono stati sottoposti  
alla valutazione di due Peer Reviewers in forma rigorosamente anonima*

© 2020 EDUCatt - Ente per il Diritto allo Studio universitario dell'Università Cattolica  
Largo Gemelli 1, 20123 Milano | tel. 02.7234.2235 | fax 02.80.53.215  
e-mail: editoriale.dsu@educatt.it (*produzione*); librario.dsu@educatt.it (*distribuzione*)  
web: www.educatt.it/libri

*Redazione della Rivista:* redazione.all@unicatt.it | *web:* www.analisinguisticaeletteraria.eu

Questo volume è stato stampato nel mese di dicembre 2020  
presso la Litografia Solari - Peschiera Borromeo (Milano)

## INDICE

Un glossario greco-latino inedito (Paris, Bibliothèque Nationale, Lat. 18556) <i>Alessandro Re</i>	5
The mediatization of femicide: a corpus-based study on the representation of gendered violence in Italian media <i>Lucia Busso, Claudia Roberta Combei, Ottavia Tordini</i>	29
Cortesía e amorevolezza nei <i>Promessi Sposi 1840</i> <i>Annick Paternoster</i>	49
“Shakespeare ist mir noch lieber”: Schumann’s Bard in the <i>Novellette</i> op. 21 n. 3 <i>Enrico Reggiani</i>	69
<i>As Painting Poetry Shall [Not] Be: An Intellectual History of Word and Image Rivalry in British Culture</i> <i>Paola Spinozzi</i>	87
Les mêmes et pourtant autres, « comme reviennent les choses dans la vie ». Odette de Crécy : prototype du personnage sériel proustien <i>Ilaria Vidotto</i>	105

## RECENSIONI E RASSEGNE

Recensioni	121
Rassegna di Linguistica generale e di Glottodidattica a cura di Giovanni Gobber	127
Rassegna di Linguistica francese a cura di Enrica Galazzi e Michela Murano	135
Rassegna di Linguistica inglese a cura di Maria Luisa Maggioni e Amanda C. Murphy	143
Rassegna di Linguistica russa a cura di Anna Bonola e Valentina Nosedà	153
Rassegna di Linguistica tedesca a cura di Federica Missaglia	159
Indice degli Autori	165
Indice dei Revisori	167



THE MEDIATIZATION OF FEMICIDE:  
A CORPUS-BASED STUDY ON THE REPRESENTATION  
OF GENDERED VIOLENCE IN ITALIAN MEDIA<sup>1</sup>

LUCIA BUSO, CLAUDIA ROBERTA COMBEI, OTTAVIA TORDINI  
ASTON UNIVERSITY (UK), UNIVERSITÀ DI BOLOGNA, UNIVERSITÀ DI PISA

This methodological study deploys hybrid techniques to investigate how femicide is framed in media. Results are consistent with ISTAT data and with the literature, and also offer novel insights. We find a tendency of not holding offenders accountable; that most femicides are perpetrated by men that victims know well; and that mediatic discourse around such crimes increases in certain circumstances and moments of the year. The analysis of the docu-fiction *Amore Criminale* reveals that metaphors are frequently used to sketch the participants' socio-psychological portraits. Iconic speech and gestures are frequently employed by interviewees to report and mime episodes of violence.

Questo studio propone un metodo ibrido per indagare la rappresentazione linguistica del femminicidio nei media italiani. I risultati sono coerenti con i dati ISTAT e con la letteratura, e offrono nuovi spunti di riflessione. Si riscontra: una tendenza a deresponsabilizzare i colpevoli; che la maggior parte dei delitti sono compiuti da uomini vicini a esse; e che su tali delitti i media si concentrano in specifiche circostanze e momenti dell'anno. L'analisi sulla *docu-fiction Amore Criminale* rivela che per delineare ritratti sociopsicologici di vittime e carnefici si impiegano metafore, mentre per descrivere/mimare episodi di violenza si impiegano strategie iconiche.

*Keywords:* corpus linguistics, multimodal analysis, Structural Topic Model, television language, journalistic language

### 1. Introduction

The social, cultural, and political issue of gendered violence has been at the centre of a heated debate in recent years, thanks to global social campaigns (such as #METOO, or *ni una menos*) that have fostered a greater attention to the problem. Nevertheless, quantitative data on the phenomenon are still scarce. According to the European Institute for Gender Equality (EIGE), no European level statistical analysis has been possible until recent years, due to complete absence of useful data (EIGE, 2018). In Italy, few – and very recent – statistics have been employed (by ISTAT and by the Department for Equal Opportunities). Findings show that family contexts are the highest-risk situations for women: in 2014 77%

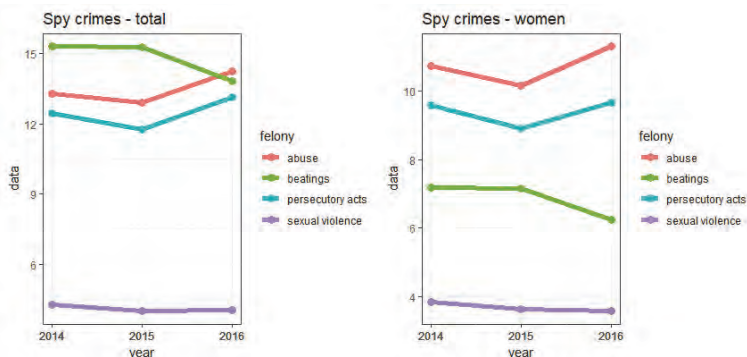
---

<sup>1</sup> Although this paper is the product of joint research, L. Busso is responsible for Sections 1 and 3, C.R. Combei is responsible for Sections 2 and 4, and O. Tordini is responsible for Sections 5 and 6.



of female murder victims were in fact killed at the hands of a partner, a former partner or a family member<sup>2</sup>. Data from the 2018 survey by SDGS (Sustainable Development Goals) on so-called spy crimes report that in 2014-2016 almost 80% of family maltreatment, 74% of persecutory acts and 89% of sexual violence concern a woman<sup>3</sup> (see Figure 1).

Figure 1 - Data on spy felonies in Italy between 2014-2016. Source: SDGS report (2018)



A new attention to the problem is also demonstrated by recent advances in the regulatory field: Law 23/4/2009, art. 612 bis c.p. introduces for the first time the crime of “persecutory acts” (stalking) in the Italian legal system<sup>4</sup>.

Notwithstanding the seriousness of such crimes and the recent global attention to the issue, Italian media and social debate still fail to truly grasp the extent of the problem. The same term ‘femicide’ is only defined for the first time in Russell<sup>5</sup> as the murder of a woman by a man because of her gender. In Italy, the word *femminicidio* is present as a literary term since the XIX century, but its current meaning and political connotations have not appeared until 2006<sup>6</sup>. The term is first introduced in the scientific literature in 2008<sup>7</sup> and is recorded by dictionaries the following year. In particular, *femminicidio* is defined in Italian as “any form of violence exercised systematically against women, in the name of an ideological patriarchal superstructure, in order to perpetuate their subordination and annihilate their identity through physical or psychological subjugation, up to slavery or death<sup>8</sup>”. Even though there is indeed an increasing number of linguistic and sociological studies focusing

<sup>2</sup> Source: Eures, III – Rapporto su Caratteristiche, dinamiche e profili di rischio del femminicidio in Italia (2016).

<sup>3</sup> Report SDGS 2018 – Informazioni statistiche per l’agenda 2030 in Italia.

<sup>4</sup> Law proposal 1455 reinforces the procedural safeguards of victims of sexual and domestic violence has recently been approved (April 2019). To date, there is also a penalty of up to six years and a € 15.000 fine for revenge porn.

<sup>5</sup> D. Russell – J. Radford, *Femicide: The Politics of woman killing*, Twayne Pub., New York 1992.

<sup>6</sup> D. Bandelli, *Femicide, gender and violence: discourse and counterdiscourses in Italy*, Palgrave Macmillan, Cham 2017.

<sup>7</sup> B. Spinelli, *Femminicidio: dalla denuncia sociale al riconoscimento giuridico internazionale*, Franco Angeli, Milano 2008.

<sup>8</sup> <https://archive.org/details/devotooli2009vocabolariodellitalianocontemporaneo> (last accessed October 12, 2020).

on the language of femicide<sup>9</sup>, most of the debate is anecdotal and discursive in nature<sup>10</sup>. We hence propose a study that aims to contribute to the social debate on gendered violence through a rigorous linguistic analysis of traditional mass-media (newspaper and television). Specifically, we propose a methodology that triangulates insights from computational and corpus linguistics, and multimodal analysis. Such an approach combines the dimensions of macro-analysis (with the uncovering of latent themes through topic modelling techniques and through corpus queries), with micro-analysis (keyword and collocational analysis, and cross-modal analysis to compare journalistic with television language).

We consider the mediatic language the perfect testbed for this methodological study, as language of the media is known to influence beliefs and sway public opinion. Our vision of the world is in fact deeply affected not only by what is reported, but also by the way in which is reported<sup>11</sup>. This process is often guided by cultural or personal stereotypes and biases. Moreover, newspapers have to give prominence to topics and writing styles which are ‘newsworthy’<sup>12</sup>, deciding what to foreground and what to background based on ‘what sells’ the most. The repetition of narrative frames and cognitive *schemata* ultimately can affect the audience’s conceptualization of events<sup>13</sup>.

Finally, this study aligns with a series of recent studies in linguistics and social sciences which strive to contextualize the mediatic representation of femicide in Italian media and culture<sup>14</sup>.

## 2. The study

Lexical analyses on corpora allow large-scale explorations which can disclose linguistic schemata and traits, either general or idiosyncratic, all of which having far-reaching societal, cultural, and cognitive implications<sup>15</sup>. In recent years, corpora have been largely used in linguistics to explore gender issues<sup>16</sup>, but to the best of our knowledge, only few studies

<sup>9</sup> F. Formato, *Gender, discourse and ideology in Italian*, Palgrave Macmillan, London 2019.

<sup>10</sup> As an anonymous reviewer correctly points out, the Accademia della Crusca has recently investigated the subject, see <https://accademiadellacrusca.it/it/consulenza/femminicidio-i-perche-di-una-parola/803> (last accessed October 27, 2020).

<sup>11</sup> N. Fairclough, *Media discourse*, Hodder Education, London 1995.

<sup>12</sup> U. Tabbert, *Language and Crime: Constructing offenders and victims in newspaper reports*, Springer, New York 2016.

<sup>13</sup> N. Fairclough, *Language and Power*, Longman, London 1989.

<sup>14</sup> S. Abis – P. Orrù, *Il femminicidio nella stampa italiana: un’indagine linguistica*, “Gender/sexuality/Italy”, 3, 2016, pp. 18-33; C. Meluzzi – E. Valvason – E. Pinelli – G. Inglese – C. Zanchi, *Violenza di genere e attribuzione dell’agentività: un’indagine tra corpus linguistics e analisi percettiva*, presented at SLI Conference, 2018.

<sup>15</sup> G. Fragaki – D. Goutsos, *Women and Men Talking About Men and Women in Greek*, in *Yearbook of Corpus Linguistics and Pragmatics 2015. Current approaches to Discourse and Translation Studies*, J. Romero-Trillo ed., Springer, New York 2015, pp. 89-115.

<sup>16</sup> P. Baker, *Using Corpora to Analyze Gender*, Bloomsbury, London 2014; L. Busso – G. Vignozzi, *Gender Stereotypes in Film Language: A Corpus-assisted Analysis*, in *Proceedings of CLiC-it 2017*, R. Basili et al. ed., Roma 2017, pp. 71-76.



focused on gendered-violence<sup>17</sup> and even fewer on how gendered-violence is framed in Italian media<sup>18</sup>.

This methodological study aims at bringing a contribution to the general understanding of the linguistic devices used to depict femicide and other gender-targeted crimes in Italian media. As mentioned in §1, we propose an assorted methodology which we claim to be advantageous in addressing this complex issue. In particular, after a short description of our data (in §2.1), in the first part of this study (in §3) we adopt a corpus-based approach, following the distinction originally made by Tognini-Bonelli<sup>19</sup> and further extended in Biber<sup>20</sup>. That is, our research employs corpus exploration as a tool to analyse previously determined systematic patterns of variation and use of linguistic features, rather than allow linguistic constructs to simply emerge from the data. Besides frequency analyses on our corpus, in §4 we apply state-of-art computational techniques to uncover recurring semantic themes in crime articles. In §5 we explore the metaphorical content of speech acts and gestures in a popular true crime television show. Concluding remarks are presented in §6.

## 2.1 Data

For all we know, several corpora of traditional or new media discourse on gendered violence have been collected so far<sup>21</sup>, but none is freely available online. Due to the scarcity of these data in open-access, we based our analyses on a specialized corpus of journalistic language (WItNECS – Women in Italian Newspaper Crime Sections) (see Table 1) and a multimedia and multimodal database of the Italian television series *Amore Criminale* (AC) (see Table 2) that we had collected as part of our previous research on this topic<sup>22</sup>. We will briefly describe these two resources for Italian below.

<sup>17</sup> S. Ehrlich, *Linguistic Discrimination and Violence against Women: Discursive Practices and Material Effects*, in *Language and Woman's Place: Text and Commentaries*, M. Bucholtz ed., Oxford University Press, Oxford 2004, pp. 223-228. J. Santaemilia-Ruiz – S. Maruenda-Bataller, *The linguistic representation of gender violence in (written) media discourse: The term 'woman' in Spanish contemporary newspapers*, "Journal of Language Aggression and Conflict", 2, 2014, 2, pp. 249-273.

<sup>18</sup> S. Abis – P. Orrù, *Il femminicidio nella stampa italiana: un'indagine linguistica*; F. Formato, *Gender, discourse and ideology in Italian*.

<sup>19</sup> E. Tognini – Bonelli, *Corpus linguistics at work*, J. Benjamins, Amsterdam 2001, Vol. 6.

<sup>20</sup> D. Biber, *Corpus-based and corpus-driven analyses of language variation and use*, in *The Oxford Handbook of Linguistic Analysis*, B. Heine & H. Narrog ed., Oxford University Press, Oxford 2010, pp. 159-191.

<sup>21</sup> K.L. Anderson – J. Cermele, *Public/Private language aggression against women*, "Journal of Language Aggression and Conflict", 2, 2014, 2, pp. 274-293; H. Purohit – T. Banerjee – A. Hampton – V.L. Shalin – N. Bhandutia – A.P. Sheth, *Gender-based violence in 140 characters or fewer: a #BigData case study of Twitter*, "First Monday", 21, 2016, 1, 6148.

<sup>22</sup> L. Busso – C.R. Combei – O. Tordini, *La rappresentazione lessicale della violenza di genere: 'Donne come vittime' nei media italiani*, in *Parola. Una nozione unica per una ricerca multidisciplinare*, B. Aldinucci et al. ed., Edizioni Università per Stranieri di Siena, Siena 2019, pp. 261-279.

Table 1 - *WItNECS statistics*

WItNECS	
Tokens: 271,506	Language: Italian
Words: 234,178	Timeframe: 13/9/2016-30/6/2017
Sentences: 10,556	Newspapers: CS, LS, R, IFQ, RRM, RF, RRN, RRP
Articles: 597	

Table 2 - *AC statistics*

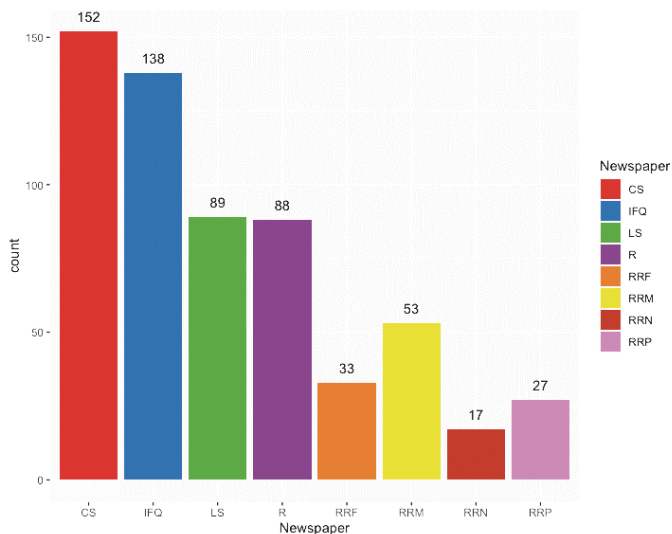
AC	
Data: videos, orthographic transcriptions, gesture and language annotations	Language: Italian
Content: docu-fiction, interviews, court records	Timeframe: 2015/2016 and 2016/2017 (2 seasons, 8 episodes)

WItNECS is a corpus of newspaper articles that was manually collected from September 13<sup>th</sup>, 2016 (the date symbolically coincides with the suicide of Tiziana Cantone, an Italian victim of revenge porn) through June 30<sup>th</sup>, 2017. For balancedness and representativity purposes, four national newspapers (online editions and archives) were considered: *Corriere della Sera* (CS), *La Stampa* (LS), *Il Fatto Quotidiano* (IFQ), and *La Repubblica* (R). Regional editions of *La Repubblica* from Milan (RRM), Florence (RRF), Naples (RRN), and Palermo (RRP) were included as well, to render the data geographically comparable.

To search for articles that focused on violence against women, several keywords and syntagmata were used<sup>23</sup>: *femminicidio* [femicide], *donna aggredita* [attacked woman], *donna deturpata* [disfigured woman], *donna maltrattata* [abused woman], *donna minacciata* [threatened woman], *donna picchiata* [beaten woman], *donna stuprata/violentata* [raped woman], *donna uccisa* [killed woman], *abuso/molestia sessuale* [sexual abuse], *delitto passionale* [crime of passion], *schiavitù sessuale* [sexual slavery], *violenza contro le donne* [violence against women], *violenza di genere* [gendered violence], *violenza sessuale* [sexual assault]. All articles that corresponded to this query were thoroughly read and only those that were relevant were included in the corpus.

As shown in Figure 2, the distribution of the 597 articles was rather homogeneous among the newspapers considered (152 articles from CS, 138 from IFQ, 218 from national and regional editions of *La Repubblica*, 89 from LS). This indicates the balancedness and representativity of the corpus, essential for our qualitative and quantitative analyses.

<sup>23</sup> Both singular and plural forms of words and syntagmata were used for the query. For the sake of brevity only singular forms are listed here.

Figure 2 - *The distribution of the articles in the corpus*

For the multimedia database, the data were collected from the true-crime television show *Amore Criminale*. The show focuses on various aspects related to violence against women in Italy, but it is also rich in terms of communicative styles and linguistic varieties and registers. In fact, each episode contains interviews with real people (e.g. victims and their family, police, psychologists, etc.), reconstructions of specific events (i.e. docu-fictions), and readings of court and legal records. The database consists of 8 episodes aired during the 2015/2016 and 2016/2017 seasons, as well the orthographic transcription and the multilevel annotation of gestures (see also §5).

### 3. Corpus-based exploration: between micro and macro analysis

Corpus linguistics tools such as the ones provided by the most common software can be extremely useful for exploring different types of patterns and phenomena, both at micro and macro level of analysis. In other words, using corpus-based methods such as keyness, collocation and concordance analysis can help shed light on issues both at linguistic and pragmatic level. We provide evidence for this claim below.

Previous work on the same corpus<sup>24</sup> revealed that at a lexical level two major semantic domains emerge in WItNECS: home (with keywords such as (*ex*) *moglie* [(*ex*) wife], (*ex*) *marito* [(*ex*) husband], *violenza domestica* [domestic violence]) and judicial field (with keywords such as *arrestare* [to arrest], *custodia cautelare* [pre-trial detention], *giudice* [judge]). Moreover, an analysis of figurative language reveals unbalanced metaphors in the representation of men and women: while women are variously represented as ‘prey’ and ‘hunted

<sup>24</sup> L. Busso – C.R. Combei – O. Tordini, *La rappresentazione lessicale della violenza di genere*, p. 273.

animals, men are never pictured as the 'hunter'. On the contrary, the use of metaphors appears to externalize the act of violence by representing it as an 'illness', or an external force. Men are often depicted as 'beasts', unable to control their instincts<sup>25</sup>.

We build on our previous findings and extend them with further analyses at both syntactical and collocational level. The former dimension was investigated via the web-based software SketchEngine<sup>26</sup>, while the latter was explored with the software package Lancs-Box<sup>27</sup>. As mentioned, women are often characterised as preys, which suggests a passive role. Moreover, there is extensive evidence in the literature that violence tends to be socially constructed as a dichotomy between an active pole (the offender) and a passive victim (the woman), and linguistically by Agent-less and passive constructions<sup>28</sup>. Therefore, we retrieved via CQL (Corpus Query Language) query all passive constructions in the corpus<sup>29</sup> and explored the arguments and adjuncts that most frequently appeared. Some interesting findings can be highlighted. Firstly, the most frequent content lemma left of the passive construction is *donna* [woman]. This suggests that women overall appear as subjects of passive constructions. One could expect that – similarly – the first words on the right would show the perpetrator of the crime. However, this is not the case. The first word to appear on the right of the passive is *a* [at], and not *da* [by], which corroborates our finding that men seem to be relieved of responsibility. We refined the query to specifically retrieve the Agent of the action (passive + *da* + *un*<sup>30</sup>), but the lemma 'man' (or husband, partner etc.) does never appear. Instead, the most frequent lemma is *raptus* [fit, burst]. In the top ten most frequent words, we also find *parente* [relative] and *extracomunitario* [illegal immigrant], the only two words referring to a person. The syntactic analysis of passive constructions hence corroborates findings on a removal of responsibility via linguistic expedients, such as different metaphorical frames or Agent-less syntactic constructions. For the collocational analysis, we will present as a case-study the behaviour of the lemmas *donna* [woman], *moglie* [wife] and *amore* [love]. We chose these three keywords for their relevance in our corpus: in fact, as already mentioned, gendered violence mainly occurs in a domestic environment, often at the hand of a (ex) husband or lover. Figure 3 graphically displays the collocation networks of our key terms<sup>31</sup> with the GraphColl tool of the Lancs-

<sup>25</sup> *Ibidem*.

<sup>26</sup> A. Kilgarriff – V. Baisa – J. Bušta – M. Jakubíček – V. Kovář – J. Michelfeit – V. Suchomel, *The Sketch Engine: Ten years on*, "Lexicography", 1, 2014, 1, pp. 7-36.

<sup>27</sup> V. Brezina – M. Timperley – T. McEnery, #LancsBox v. 4.x [software], 2018. Available at: <http://corpora.lancs.ac.uk/lancsbox> (last accessed October 27, 2020).

<sup>28</sup> G. Bohner, *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, 2001, pp. 515-529; D. Bandelli, *Femicide, gender and violence: discourse and counterdiscourses in Italy*; C. Meluzzi – E. Valvason – E. Pinelli – G. Inglese – C. Zanchi, *Violenza di genere e attribuzione dell'agentività*.

<sup>29</sup> CQL query in SketchEngine: [lemma="essere|andare|venire"] [tag="V.\*:ppast"] ?[tag="V.\*:ppast"].

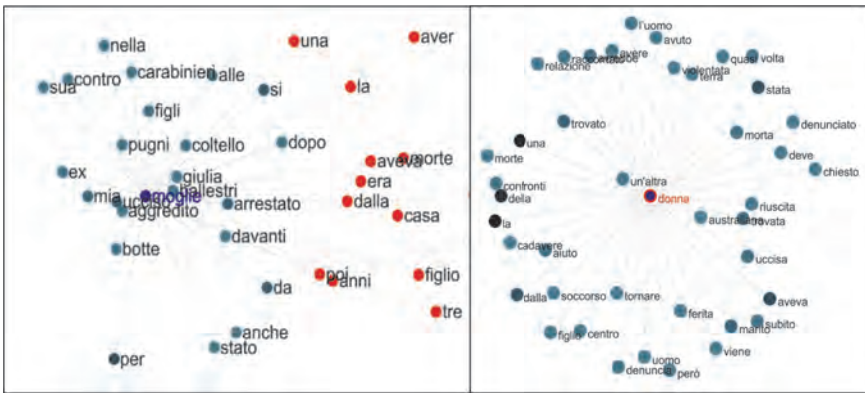
<sup>30</sup> [lemma="essere|andare|venire"] [tag="V.\*:ppast"] ?[tag="V.\*:ppast"] [lemma="da"] [lemma="un"].

<sup>31</sup> M. Phillips, *Aspects of text structure: An investigation of the lexical organisation of text*, North-Holland, Amsterdam 1985.

box software. The tool identifies collocations and displays them as a collocation graph or network using different statistical measures of association.

The graph displays three dimensions: i) strength of collocation, which is indicated by the distance between the node and the collocates ('magnet effect'); ii) collocation frequency, which is indicated by the intensity of the colour of the collocate and iii) position of collocates, which reflects the exact position (left or right) of the collocates in text. For the present work, we chose a span of 3 collocates to the right and to the left, and to use Mutual Information (MI) as statistical association measure. A threshold MI of 3 was selected as well to avoid over-populated graphs.

Figure 3 - Collocation network of the terms *donna* and *moglie*



From this preliminary exploration, some interesting considerations already surface. As we can see from Figure 3, female victims (either women or wives) have both collocates which point to the abuse (e.g.: *pugni* [punches]; *coltello* [knife]; *aggredata* [assaulted]; *uccisa* [killed]). Furthermore, *donna* and *moglie* share a considerable number of collocates, indicating that they are often used in the same contexts: women are often wives. Interestingly, the shared collocates concern not only the crime (*morte* [death]) but mainly the family domain (*casa* [home]; *figlio* [son/child]). In general, women are associated with the family nucleus: they are mothers (*madri*) of children (*bambini*). Notably, collocations of *donna* include also a few past participles (*uccisa* [killed]; *ferita* [harmed]; *trovata* [found]), which are consistent with the finding of women being portrayed in passive roles. The two terms mostly co-occur with other nouns and verbs, adjectival modifiers are extremely rare, as journalistic language is pre-eminently denotative<sup>32</sup>.

The graph in Figure 4 visually depicts the collocates of the lemma *amore*, which in Italian is often used as a metonymy for relationship. Since most episodes of violence occur between (ex) partners, in fact, we considered it an interesting linguistic cue to explore how newspaper address the issue of violent relationships, a widely spread sociological phenomenon which oftentimes result in abuse, stalking or murder. The lexical collocates reveal that

<sup>32</sup> L. Busso – C.R. Combei – O. Tordini, *La rappresentazione lessicale della violenza di genere*.

violent relationships are frequently defined metaphorically as *amore malato* [sick love]. The analogy between love and illness is extremely well-known<sup>33</sup>. However, its use in the context of journalistic language describing murder cases is at least controversial. By equating a violent relationship to an illness, the perpetrator of the violence is not held completely accountable for his actions, as violence is merely a symptom of an uncontrollable disease. Violence is further defined as *non amore* [non-love], and *finito* [ended]. In this case as well the responsibility is shifted from the perpetrator to the overall cause of an ended relationship or an unreciprocated love.

In sum, a corpus-driven exploration of the data allows for a preliminary understanding of phenomena at both a macro-level and a micro-level of analysis, such as (respectively) the occurrence of Agents and Patients in passive constructions in the corpus, and the specific collocates. We further expand the analysis by employing computational techniques (macro-level) and by crossmodally comparing journalistic language to the language of TV (micro-level).

Figure 4 - Collocation network of the lemma *amore* [love]



#### 4. Topic analysis in *WIINECS*

In recent years, as a result of the ever-growing interest and availability of textual material, various machine-assisted approaches for the analysis of corpora have been proposed. Topic modelling is one of the most accessible and widely used techniques as it can inductively identify semantic themes in large corpora, even when data are unstructured<sup>34</sup>.

This study employs a special kind of topic models, namely the Structural Topic Model<sup>35</sup>, to explore the issues on which newspapers concentrate when they depict violence against

<sup>33</sup> Z. Kövecses, *The Language of Love*, Bucknell University Press, London 1988.

<sup>34</sup> A. Murakami – P. Thompson – S. Hunston – D. Vajn, 'What is this corpus about?: using topic modelling to explore a specialised corpus', *Corpora*, 12, 2017, 2, pp. 243-277; C.R. Combei – D. Giannetti, *The Immigration Issue on Twitter Political Communication. Italy 2018-2019*, *Comunicazione Politica*, 21, 2020, 2, pp. 231-263.

<sup>35</sup> M.E. Roberts – B.M. Stewart – D. Tingley – C. Lucas – J. Leder-Luis – S.K. Gadarian – D.G. Rand, *Structural Topic Models for Open-Ended Survey Responses*, *American Journal of Political Science*, 58, 2014, 4, pp. 1064-1082.



women. Moreover, this approach enables the observation of the narrative variation and evolution over the timeframe considered and across the newspapers included in the corpus.

#### 4.1 Structural Topic Model

In the fashion of other frameworks of topic modelling, such as Latent Dirichlet Allocation<sup>36</sup>, the Structural Topic Model (STM<sup>37</sup>) is a generative model of word counts made available for R<sup>38</sup> through the *stm* package<sup>39</sup>. Particularly, STM is an unsupervised learning method that uses modelling assumptions, text properties, and covariates to estimate topics within a corpus and to organize its documents on the basis of co-occurrences of words. In the STM framework, a topic is a mixture over words where each word has a probability of belonging to a topic, while a document is a mixture over topics.

The major advantage of STM is that the prior distribution of topics may be varied as a function of covariates. A covariate is a variable that has a statistical relationship with the dependent variable; that is, a covariate can explain or predict a dependent variable. Covariates are used in topic modelling to examine relationships among variables in a regression-like scheme and to further explain topical prevalence or topical content. This feature is also particularly relevant for the validation and interpretation of topics in corpora. The STM has been greatly employed in social and political science research<sup>40</sup>, but to the best of our knowledge, it is still underexploited in linguistics<sup>41</sup>.

#### 4.2 Text pre-processing

Since all newspaper articles as well as their metadata would have been employed in topic modelling experiments, various pre-processing steps were performed on the raw data. We defined an *ad-hoc* ‘stop words’ list for Italian consisting of 1,300 lexically empty or uninformative words (i.e. prepositions, conjunctions, determiners, auxiliary verbs, etc.). Next, we cleaned and normalized the corpus with several functions available on these R pack-

<sup>36</sup> D.M. Blei – A.Y. Ng – M.I. Jordan, *Latent Dirichlet Allocation*, “Journal of Machine Learning Research”, 3, 2003, 3, pp. 993-1022.

<sup>37</sup> Throughout the paper we will use “STM” to abbreviate the Structural Topic Model framework and “stm” to refer to the R package (M.E. Roberts et al., *Structural Topic Models for Open-Ended Survey Responses*).

<sup>38</sup> R Core Team, *R: A language and environment for statistical computing*, [computer program], available at <https://www.r-project.org>. R Foundation for Statistical Computing, Wien 2018.

<sup>39</sup> M.E. Roberts et al., *Structural Topic Models for Open-Ended Survey Responses*.

<sup>40</sup> L. Pinto, *Structural Topic Model per le scienze sociali e politiche*, “POLIS”, 33, 2019, 1, pp. 163-174.

<sup>41</sup> G. Brookes – T. McEnery, *The utility of topic modelling for discourse studies: A critical evaluation*, “Discourse Studies”, 21, 2019, 1, pp. 3-21; C.R. Combei, *Is ‘manovra’ really ‘del popolo’? Linguistic Insights into Twitter Reactions to the Annual Italian Budget Law*, in *Proceedings of the Sixth Italian Conference on Computational Linguistics (CLiC-it 2019)*, R. Bernardi et al. ed., CEUR-WS, Aachen/Bari 2019, pp. 1-6.

ages: `quanteda`<sup>42</sup>, `tm`<sup>43</sup>, and `qdapRegex`<sup>44</sup>. In particular, all letters were converted to lower case; numbers, punctuation, sequences made up of one character, and excessive white spaces were deleted. In order to use the metadata of the newspaper articles (i.e. the name of the newspaper and the publication date of the article) as covariates for topical prevalence, this information was coded in the file names during data collection and further extracted and processed with R (e.g. from ‘CS\_2016\_12\_14.txt’ we obtained the newspaper name *Corriere della Sera* and the date ‘2016-12-14’).

### 4.3 Topics

Starting from the articles in WItNECS and using the `textProcessor` and `prepDocuments` functions on the `stm` package, we generated a corpus containing 597 documents, 77,087 single word forms, 17,840 lemmas, and the metadata of the articles. After having applied the `stm searchK` function to perform held-out likelihood, semantic coherence, and residual analysis tests, the number of topics was set to 20 (this model had high levels of held-out likelihood and semantic coherence as well as a low level of residuals). STM also allowed us to choose the type of initialization; here we used the spectral initialization, as previous studies proved its stability and consistence<sup>45</sup>. As a word profile we chose the highest probability. Considering that the stemming options available through R did not perform well on our corpus of Italian texts and that we were interested in the gender of the words, for our STM model we used word forms instead of stemmed words.

Figure 5 shows the 20 topics that emerge from the STM model on WItNECS together with the macro-topic labels we assigned to them. The labels are based on the examination of the most semantically coherent words for each topic and on our reading of the most exemplar documents (obtained with the `findThoughts` function on `stm`), namely newspaper articles that have the highest proportion of words associated with each topic.

The results in Figure 5 display interesting insights. First of all, judging by the topic proportions and the word probabilities of the STM model, it seems that the most prevalent topics in WItNECS are domestic violence and uxoricides (topics: 1, 4, 6, 14, 16, 18), with an overall expected proportion of 32%. Some of the words that have high values of probability in the macro-topic of domestic violence and uxoricides are *marito* [husband], *fidanzato* [fiancé], *relazione* [relationship], *coppia* [couple]. This preliminary finding suggests that domestic violence is the most debated issue in our corpus. Articles presenting cases of sexual harassment are also rather recurrent as suggested by the high proportion of

<sup>42</sup> K. Benoit – K. Watanabe – H. Wang – P. Nulty – A. Obeng – S. Müller – A. Matsuo, *quanteda: An R package for the quantitative analysis of textual data*, “Journal of Open Source Software”, 3, 2018, 30, pp. 774-778.

<sup>43</sup> I. Feinerer – K. Hornik – D. Meyer, *Text Mining Infrastructure in R*, “Journal of Statistical Software”, 25, 2008, 5, pp. 1-54.

<sup>44</sup> T.W. Rinker, *qdapRegex: Regular Expression Removal, Extraction, and Replacement Tools*. 0.7.2., University at Buffalo, Buffalo, New York 2017.

<sup>45</sup> M.E. Roberts – M. Stewart Brandon – D. Tingley, *Navigating the Local Modes of Big Data: The Case of Topic Models*, in *Computational Social Science: Discovery and Prediction (Analytical Methods for Social Research)*, R.M. Alvarez ed., Cambridge University Press, Cambridge 2016, pp. 51-97.

topics (i.e. 21%) related to this issue (topics: 5, 7, 9, 13, 20). Some of the terms strongly associated with this macro-topic are: *aggressione* [aggression], *sessuali* [sexual], *molestie* [assault], and *abusi* [abuses]. Next, topics 8 and 17 have an expected topic proportion of 13% and they reflect how newspapers delineate crime scenes, giving deictic information on the crime setting and its participants, and using words and syntagma that pertain to the semantic field of crime scenes (e.g. *corpo* [body], *cadavere* [corpse], *appartamento* [apartment], *fiamme* [flames], *coltello* [knife], etc.). Similarly, topics 3 and 19 (expected topic proportion: 12%) concern news on arrests, accusations and sentences for femicides or other types of gender-based crimes; thus, most words in these topics belong to the semantic field of law (e.g. *gip* [magistrate in charge of preliminary investigations], *custodia cautelare* [pre-trial detention], *corte* [court], *pena* [punishment], *condanna* [sentence], etc.). Other topics (i.e. 2 and 12) that emerge from the model regard initiatives of helping victims and preventing femicides (estimated topic proportion: 7%), as suggested by the words that have the highest probability of belonging to this topic: *accoglienza* [community shelter], *centri* [centres], and *antiviolenza* [fighting against violence]. Topics 11 and 15 (estimated topic proportion: 9%) reflect the newspaper attention for new media and revenge porn (e.g. *video hard*, [pornographic video], *social* [social network], *facebook*, *diffusione* [distribution], *rete* [network], etc.). Finally, topic 10 (estimated topic proportion: 6%) regard the treatment and the care provided to victims (e.g. *ospedale* [hospital], *progognsi* [prognosis], *lesioni* [injuries], *soccorso* [rescue], *ricoverata* [hospitalised], etc.).

Figure 5 - Topics in *WItNECS*

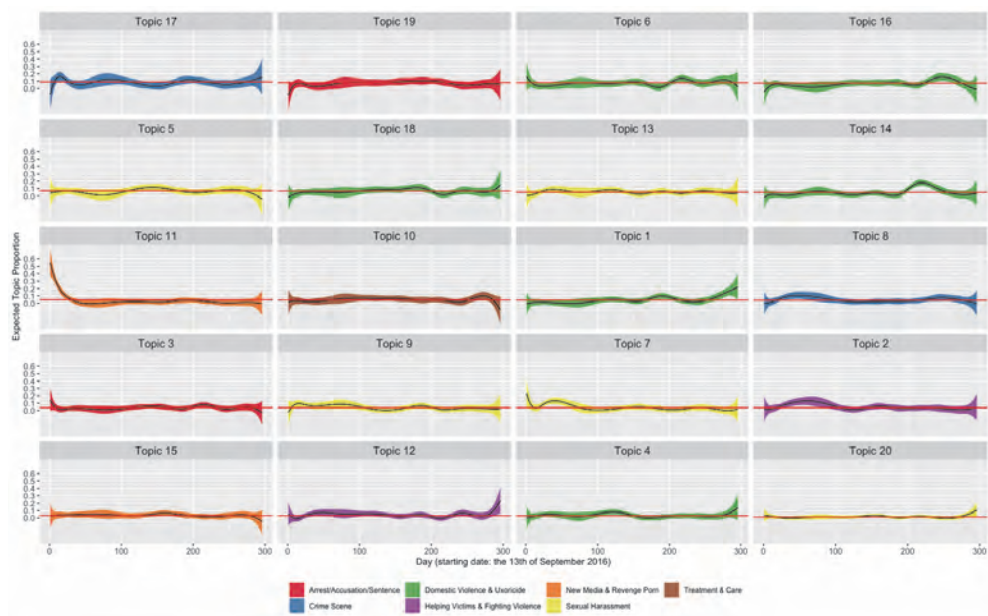
Since the other aim of this study was to reveal temporal patterns as well as newspaper-dependent communicative intentions, we also extracted and plotted the estimated effects for two covariates (i.e. 'date', 'newspaper'). The gamma matrix and the 'findThoughts' function on stm guided us in the examination of the distribution of topics throughout our 9-months timeframe and across the newspapers considered. At the same time, these features allowed us to assess the goodness of our model. Several tendencies are traceable, but we will only mention results that reached statistical significance (p-values < 0.05 in the regression-like scheme).

In Figure 6, the variation of media attention on specific topics is displayed. On the  $x$  axis, the days are counted, starting from September 13<sup>th</sup>, 2016, while on the  $y$  axis the expected topic proportions are shown. First, it seems that newspapers wrote about initiatives of fighting femicides and helping female victims of violence mainly in November, although there is a rising trend also in February and March. This is not surprising, as these peaks correspond to the International Day for the Elimination of Violence against Women, Valentine's Day, and the International Women's Day and the immediate days before or after these events. Next, topics on the use of new media for the distribution of revenge porn are mostly prevalent in conjunction with the aftermath of Tiziana Cantone's suicide and the related succession of events in time (e.g. funeral, court trials, etc.), even if there are some other news that focuses on similar issues (e.g. various cases of blackmails and revenge porn involving young women, some of them under-aged). The topic variation over time also suggests that generally news regarding sexual harassment were more frequent at the beginning (October/November) and in the middle (February/March) of the timeframe considered in this study. Similar trends are observable for news on generic femicide and articles that describe crime scenes; apparently there are higher estimated topic proportions in October, December, and February. In Figure 5, it was clear that that most crimes recounted in our corpus were classifiable as cases of domestic violence or uxoricides. According to the data plotted in Figure 6, various articles presenting stories of violence inflicted by male partners were written in February, March, May, and June; although more in-depth analyses on more data should be performed to confirm this trend, the peaks in February and March appear to occur in correspondence with Valentine's Day and International Women's Day. Although these findings are interesting and statistically significant, as far as we can tell, some could be due to chance; as a matter of fact, research in sociology and criminology has fallen short on explaining specific temporal trends in femicide and other gendered violence<sup>46</sup>. Finally, topics on the aftermath of the crime (i.e. arrests, hearings, trials, and sentences, as well the victims' treatment and care) were constant during the entire time span covered in this study.

---

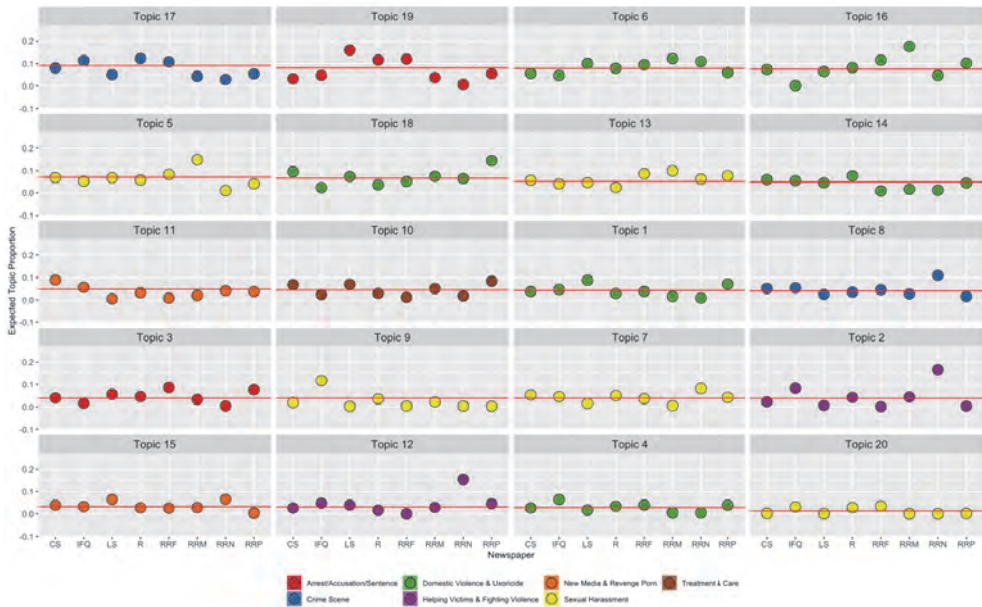
<sup>46</sup> S. Caman – M. Jristiansson – S. Granath – J. Sturup, *Trends in rates and characteristics of intimate partner homicides between 1990 and 2013*, "Journal of Criminal Justice", 49, 2017, 1, pp. 14-21; M. Dawson – M. Carigan, *Identifying femicide locally and globally: Understanding the utility and accessibility of sex/gender-related motives and indicators*, "Current Sociology", 68, 2020, 6, pp. 1-20.

Figure 6 - Variation of topics over time



As regards the distribution of topics in the newspapers examined, based on Figure 7 and the statistical analyses performed, it is rather difficult to map strong tendencies. Nevertheless, it seems that initiatives of helping victims and preventing femicides are reviewed more often by journalist from IFQ and RRN. Higher than the average levels of expected topic proportions for crimes involving new media and revenge porn are traceable in articles from RRN and CS. Next, topics concerning domestic violence have a significantly high expected proportion in RRM, RRP, and LS, suggesting that, generally, this kind of news is covered more in regional newspapers. Additionally, a higher expected proportion of topics pertaining to arrests and sentences for cases of femicide or other types of gendered violence against women seems to be ascribable to LS. Topics referring to sexual harassment are rather frequent in the IFQ articles. Words and syntagmata belonging to the topic of crime scene imagery tend to appear more often in the national and regional editions of *La Repubblica*. Finally, the RRP, LS and CS articles in WItNECS dedicate more space than others to the aftermath of the crime and to how victims are treated.



Figure 7 - *Variation of topics across newspapers*

All in all, in this section we have shown how the STM framework can be successfully employed to extract the topics on which newspapers concentrate when they discuss about violence against women. Furthermore, this technique allowed us to explore how the issue of gendered violence varied over the time period considered and across the four newspapers surveyed in WITNECS, revealing interesting linguistic and societal trends that could be confirmed in more in-depth future research.

### 5. Discourse and gestures analysis of *Amore Criminale*

*Amore Criminale* (AC) is a popular Italian TV show which has been broadcasted since 2007. It is possible to identify three main types of episodes in the show. Most frequently, a love story is narrated in which a woman gets killed by her partner or ex-partner. These episodes include interviews with the victims' families, friends and co-workers, but also with the police and lawyers involved in the trials. Other episodes comprise interviews to women who survived both physical and psychological violence. In this case, they are invited to talk about the dramatic events from their personal perspective, and to share thoughts and emotions with the interviewer and the audience. Some episodes also include interviews to men who had previously perpetrated violence on women (i.e. *uomini maltrattanti*), and who are currently being treated in specialized centers. These men are asked to report (anonymously) their experience as serial aggressors and to describe the escalation of brutality against their partners. Finally, all episodes contain docu-fictions and the readings of legal material. The presenter acts as a narrator, which recaps the events and provides further details to the story.



The choice to analyse AC lies in the fact that it offers a large and manifold insight into the dynamics of femicide and other acts of violence against women in Italy. As already outlined in § 2.1, the program is also characterised by the use of diverse communicative styles and registers. Undoubtedly, all these features allow us to obtain a wider perspective on gendered violence, and thus to better grasp its complexity.

For what concerns the analysis of the AC database (see § 2.1), we employed the ELAN software, a tool for multimodal transcription and processing<sup>47</sup>. Our aim was to provide a coherent methodological basis for the annotation and examination of audio and video resources and the interpretation of the multimedia content. ELAN allows us to create and manipulate multiple annotations on different tiers, namely multiple levels of analysis in which both linguistic and metalinguistic information can be segmented and labelled. The transcription and the analysis of multimodal data builds on Magno Caldognetto et al.'s<sup>48</sup> methodological frameworks. Based on their *partitura* system, communicative acts can be processed either separately or in their mutual interaction. So, we were able to individuate utterances and gestures of each participant – interviewer and interviewed – and to analyse them either singularly and simultaneously. Next, we applied the norms for orthographic transcription and coding proposed by Kong et al.<sup>49</sup>. This methodology was chosen as it incorporates and homogenises previous classifications, respectively proposed by Ekman and Friesen<sup>50</sup>, McNeill<sup>51</sup> and Mather<sup>52</sup> (see Poggi and Wagner et al. for a review<sup>53</sup>).

This classification method was employed to label either the gestures of the participants (presenter, actors, victims/relatives in interviews), and the content of their oral productions (for example, M-speech act). Specifically, we employed the following tags:

- ‘I-gesture’ for iconic gestures that depict visual information, shape the form of an object and/or try to reproduce a dynamic movement (i.e. the act of miming a real scene);
- ‘M-gesture’ for metaphoric gestures, which convey and abstract idea to conceptualize a concrete experience;

<sup>47</sup> Available at <https://tla.mpi.nl/tools/tla-tools/elan/> (last accessed July 10, 2019); H. Sloetjes, *ELAN: Multimedia annotation application*, in *The Oxford Handbook of Corpus Phonology*, J. Durand et al. ed., Oxford University Press, Oxford 2014, pp. 305-320.

<sup>48</sup> E. Magno Caldognetto – P. Cosi – F. Cavicchio – I. Poggi, *La ‘partitura’: un sistema di annotazione multilivello di interazioni multimodali basata su ANVIL*, in *Atti delle XIV Giornate di Studio del GFS*, A. De Dominicis et al. ed., Esagrafica, Roma 2004, pp. 253-258.

<sup>49</sup> A.P. Kong – S. Law – C.C. Kwan – C. Lai – V. Lam, *A Coding System with Independent Annotations of Gesture Forms and Functions during Verbal Communication: Development of a Database of Speech and GESTure (DoSaGE)*, “Journal of Nonverbal Behaviour”, 39, 2015, 1, pp. 93-111.

<sup>50</sup> P. Ekman – W.V. Friesen, *The repertoire of nonverbal behavior: Categories, origins, usage, and coding*, “Semitica”, 1, 1969, 1, pp. 49-98.

<sup>51</sup> D. McNeill, *Hand and mind: What gestures reveal about thought*, University of Chicago Press, Chicago 1992.

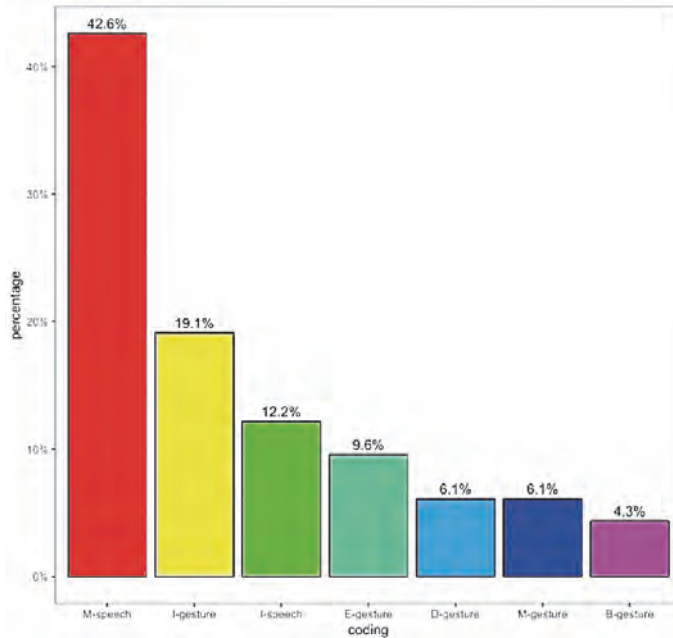
<sup>52</sup> S.M. Mather, *Ethnographic research on the use of visually based regulators for teachers and interpreters*, in *Attitudes, innuendo, and regulators*, M. Metzger – E. Fleetwood ed., Gallaudet University Press, Washington 2005, pp. 136-161.

<sup>53</sup> I. Poggi, *Mind, hands, face and body. A goal and belief view of multimodal communication*, Weidler, Berlin 2007; P. Wagner – Z. Malisz – S. Kopp, *Gesture and speech in interaction: An overview*, “Speech Communication”, 57, 2014, pp. 209-232.

- ‘D-gesture’ for deictic gestures that point to concrete locations, used by the speaker to indicate a specific element in the physical communicative context (i.e. pointing finger);
- ‘E-gesture’ for emblematic gestures, i.e. with standard/shared properties (for instance, gestures which are not language-specific, such as the thumb-index gesture for ‘OK’).

In total, we analysed about 16 hours of the TV show, and tagged 116 observations. The content and percentages resulting from these classifications are shown in the plot in Figure 8 below.

Figure 8 - Percentages resulting from the gesture-speech tagging



Results show that 48.7% of communicative acts<sup>54</sup> in the AC database have a metaphorical content. If we consider the correspondences with the audio-visual material, we note that metaphors are specifically employed to describe the behaviour of the subjects – both men and women – involved in the events. These metaphors are drawn from various semantic domains, for example: nature/animals/plants (“*era un parassita*” [he was a parasite]; “*scatta la caccia all’assassino*” [the killer hunt starts]; “*Francesca era un leone [...] una combattente*” [Francesca was a lion [...] a fighter]; “*il fuoco della gelosia*” [the jealousy fire]), perception (“*lento sprofondare nel mondo oscuro del ragazzo che amava*” [slowly sinking into the dark world of the man she loved]; “*refrattario a ogni luce di umanità*” [impervious to any light of humanity]), war (“*eravamo tutti nel suo mirino*” [we were all in his gunsight];

<sup>54</sup> For both iconic and metaphorical acts, we summed the total of speech acts and gestures (for instance, 42.6% M-speech + 6.1% M-gestures).

“*assediate da un uomo*” [besieged by a man]), or technology and games (“*piano piano [...] ho cominciato a vederla come un avversario*” [I was starting to see her as a rival]; “*era come se avessi combattuto in un videogame*” [it was like as if I fought in a videogame]; “*per me lei era solo una pedina*” [she was just a pawn to me]). Specifically, these last examples are retrieved from the episodes including interviews with *uomini maltrattanti*.

Interestingly, we note that in their interviews these men adopt recurrent schemes that might suggest a displacement of violence outside their own moral dimension. That is, they consistently report to have undergone a violent *raptus* (such as *rabbia cieca* [blind rage]) that has temporarily taken over their consciousness, making them incapable of controlling their behaviour and of acting responsibly. Moreover, most of them report to have faced past traumas as victims of violence, which left a permanent trail on their personality. At first glance, one might suppose that these men are still looking for an external justification of their viciousness. Nevertheless, from these interviews we may infer that they are deeply aware of what they have inflicted on the victims. In fact, they have autonomously decided to undergo psychological treatments in specialized rehabilitation centres.

Overall, gestures revealed to be the most significant strategy to convey the authentic content of the AC show, namely the characterization at socio-psychological level of both victims and *uomini maltrattanti*. It is also worth noting the meaningful presence of I-gestures (19.1%), combined with I-speech acts (12.2%). Iconic pragmatic strategies are employed either by victims (if any) and by the presenter/interviewer to describe or mime acts of violence, (for instance, “*avevo un occhio così*” [I had an eye like that]; “*mi acciappa per i capelli e mi tira fuori*” [he grabs me by my hair and pulls me out]). In some cases (6.1%), I-gestures are supported by D-gestures (“*ero tutta piena di lividi [...] qua [...] dappertutto*” [I was all full of bruises [...] here [...] everywhere]). Figures 9 and 10 provides two examples of multimodal annotations, both performed on interviews with survived victims:

Figure 9 - Annotation of AC in ELAN (I-speech + I-gesture)



In the first picture, the victim utters “*avevo un occhio così...*” [I had an eye like that]. She describes the concrete act of violence by shaping the form of a swollen eye – which resulted from a dramatic rebellion attempt against her partner. Therefore, the combination of speech and gesture can be here classified as iconic.

Figure 10 - *Annotation of AC in ELAN (M-speech + M-gesture)*



In Figure 10, we see a victim describing the moment immediately before the aggression. She dwells here on some details of the scene – in particular on her ex-partner’s face. Here, the woman specifically depicts “*il fuoco che gli ho visto negli occhi...*” [the fire I saw in his eyes]. Since she employed an abstract image to conceptualize a visual experience, we categorized the content of this passage as metaphoric.

Further considerations resulting from the multimodal analysis on AC will be presented in § 6, in light of the comparison with the WItNECS corpus.

## 6. Concluding remarks

In this paper, we presented an innovative methodology to analyse linguistic and pragmatic phenomena in mediatic language. Our approach triangulates quantitative and qualitative evidence from corpus linguistic and content analysis, computational modelling, and multimodal analysis. Results integrate macro-level with micro-level analysis, to shed light on the linguistic representation of femicide in contemporary Italian media.

The journalistic corpus was firstly explored at the syntactic and collocational level. This step was essential to a preliminary understanding of the corpus, as the analysis revealed patterns of recurrent usage. These patterns define the linguistic profile of the journalistic language used by the newspapers in our corpus. For instance, we find that women are often referred to in passive constructions, but that perpetrators are seldom used as Agents. Moreover, metaphoric language tends to picture violence as an extraneous force, rather

than the act of a perpetrator. Women are also found to be often described ‘in-relation-to’ their close family members.

To refine our preliminary analysis, we used the STM technique for more in-depth quantitative explorations. This allowed us to display and organize the content of 597 crime news articles. Besides revealing cues on the stylistic simplicity and objectivity of crime news, the major finding of this investigation is that most crimes are committed by someone in the victim’s domestic circle (e.g. husband, ex-boyfriend, etc.). These quantitative findings – that fully confirm data surveyed by EIGE, ISTAT and the Department for Equal Opportunities – have not been reported in other studies in linguistics and/or discourse analysis so far. Moreover, the temporal covariate in the STM model provided interesting patterns of topic variations in the time span considered: apparently, more news on initiatives of fighting femicides, on cases of sexual harassments, and on domestic violence was published in November, February, and March. That is, in correspondence with the International Day for the Elimination of Violence against Women, Valentine’s Day, and International Women’s Day. Regarding the distribution of topics in the newspapers under consideration, very few patterns were traceable: news depicting crime scenes was more recurrent in the national and regional editions of *La Repubblica*; crimes regarding the distribution of revenge porn on new media appeared more often in CS and in RRN; finally, the issue of domestic violence was frequent in LS, RRM and RRP.

The multimodal analysis on the AC database allowed us to directly compare two different mediatic discursive styles and communicative strategies. Our results shed light on a large variety of patterns, which highly depend on the focus of the interview/presentation. In both discourse and gestures, metaphors are frequently used to sketch the socio-psychological portrait of culprits and victims. Undoubtedly, this type of strategy contributes to the spectacularization of events presented in the television show, which contrasts with the descriptive approach that traditionally characterizes newspapers. Finally, iconic speech and gestures are also frequently adopted by both victims and interviewees to report, testify and/or mime episodes of violence – either physical or psychological.

In the context of the representation of femicide in Italian media, we implemented a novel methodology. Although preliminary, our findings corroborate the usefulness of such interdisciplinary approach – at the crossroads of corpus and computational linguistics, communication, and social studies. Combining quantitative methods and rigorous qualitative research practices allowed us to achieve a deeper understanding on the linguistic, social and cultural implications of gender-based violence, and shed light on latent aspects of linguistic and textual data which may not be immediately visible.



FACOLTÀ DI SCIENZE LINGUISTICHE E LETTERATURE STRANIERE  
**L'ANALISI LINGUISTICA E LETTERARIA**

ANNO XXVIII - 3/2020

EDUCatt - Ente per il Diritto allo Studio Universitario dell'Università Cattolica  
Largo Gemelli 1, 20123 Milano - tel. 02.72342235 - fax 02.80.53.215  
e-mail: editoriale.dsu@educatt.it (produzione)  
librario.dsu@educatt.it (distribuzione)  
redazione.all@unicatt.it (Redazione della Rivista)  
web: www.educatt.it/libri/all

ISSN 1122 - 1917



9 788893 357661