# Explaining Intimate Partner Violence with LLaMAntino

Pierpaolo Basile[1], Marco de Gemmis[1,*], Elio Musacchio[1], Marco Polignano[1], Giovanni Semeraro[1], Lucia Siciliani[1], Vincenzo Tamburrano[1], Vita Barletta[1], Danilo Caivano[1], Fabiana Battista[2], Antonietta Curci[2], Rosa Scardigno[2], Gabriella Calvano[3] and Patrizia Sorianello[3]

[1]University of Bari Aldo Moro, Dept. of Computer Science, Via E. Orabona 4, Bari, 70125, Italy
[2]University of Bari Aldo Moro, Dept. of Education Science, Psychology, Communication Science, Via Scipione Crisanzio 42, Bari, 70122, Italy
[3]University of Bari Aldo Moro, Dept. of Humanistic Research and Innovation, Via Scipione Crisanzio 42, Bari, 70122, Italy
[4]University of Bari Aldo Moro, Dept. of Humanistic Research and Innovation, Piazza Umberto I, Bari, 70121, Italy

**Abstract**

Violence perpetrated to their own partner is a social issue that can take place in different forms and in different settings (i.e., in person, online). These different forms of violence can be circumscribed into two broad categories known as Intimate Partner Violence (IPV) and Cyber Intimate Partner Violence (C-IPV). Social Media and technologies can exacerbate these types of behaviors but some "digital footprints", such as textual conversations, can be exploited by Artificial Intelligence models to detect and, in turn, prevent them. With this aim in mind, in this paper, we describe a scenario in which the Italian Language Model family LLAmAntino can be exploited to explain the presence of toxicity elements in conversations related to teenage relationships and then educate the interlocutor to recognize these elements in the messages received.

**Keywords**

Natural Language Processing, Abusive Language

## 1. Introduction

Studies so far have shown that one of the most common types of violence is the one committed towards their own partner, namely intimate partner violence. Due to the high rate of these behaviors in society, their early detection can be useful in reducing them. A fruitful way to reach this goal is by building AI models to discriminate against possible violence-related behaviors. Indeed, the identification of these behaviors can be problematic for victims due to the nature of the relationship with their perpetrator. In fact, people continue to hold disbelief concerning romantic engagement, which can turn into acceptance of harmful behaviors. Therefore, having a tool that can help in identifying possible violent behaviors could serve as a preventive measure for the exacerbation of harmful situations. In particular, we propose the adoption of Large Language Models (LLMs) to explain the presence of toxicity elements in a dataset of conversations related to teenage relationships. We are convinced that this novel approach, which provides the reasons why a message represents violence, can educate the interlocutors and promote partner violence prevention.

The paper is structured as follows: in Section 2, we provide a frame of what is intimate partner violence, the different forms, and the deleterious intra and interpersonal consequences.

In Section 3, we briefly describe the LLM we adopted in our scenario. Section 4 focuses on the task of explaining toxic language in the context of IPV. We describe the dataset and the different types of annotations provided by researchers in General Psychology, as well as the prompting strategy adopted to instruct the language model. Finally, in Section 5, we draw some conclusions and discuss directions for the continuation of the work.

*Ital-IA 2024: 4th National Conference on Artificial Intelligence, organized by CINI, May 29-30, 2024, Naples, Italy*
*Corresponding author.
†These authors contributed equally.
pierpaolo.basile@uniba.it (P. Basile); marco.degemmis@uniba.it (M. d. Gemmis); elio.musacchio@phd.unipi.it (E. Musacchio); marco.polignano@uniba.it (M. Polignano); giovanni.semeraro@uniba.it (G. Semeraro); lucia.siciliani@uniba.it (L. Siciliani); vincenzo.tamburrano@uniba.it (V. Tamburrano); vita.barletta@uniba.it (V. Barletta); danilo.caivano@uniba.it (D. Caivano); fabiana.battista@uniba.it (F. Battista); antonietta.curci@uniba.it (A. Curci); rosa.scardigno@uniba.it (R. Scardigno); gabriella.calvano@uniba.it (G. Calvano); patrizia.sorianello@uniba.it (P. Sorianello)
0000-0002-0545-1105 (P. Basile); 0000-0002-2007-9559 (M. d. Gemmis); 0000-0002-3939-0136 (M. Polignano); 0000-0001-6883-1853 (G. Semeraro); 0000-0002-1438-280X (L. Siciliani); 0009-0007-3802-842X (V. Tamburrano); 0000-0002-0163-6786 (V. Barletta); 0000-0001-5719-7447 (D. Caivano); 0000-0003-4086-739X (F. Battista); 0000-0002-0932-7152 (A. Curci); 0000-0002-5725-6483 (R. Scardigno); 0000-0003-2780-9902 (G. Calvano); 0000-0002-6632-0555 (P. Sorianello)

## 2. What is Intimate Partner Violence: Definition and Forms

In 2023, the World Health Organization's (WHO) report underlined an increasing rate of women's death due to intimate partner violence, almost 5% higher than the one detected in 2017. Indeed, intimate partner violence does not occur only in terms of physical violence (e.g., violence that exacerbates until victims' death) but also in other multiple forms and it is not related only to women but can be perpetrated towards men as well. Intimate partner violence has been defined as all forms of abuse and/or aggression performed by a partner to their own partner[1]. Consequently, four patterns of categories can be identified (i.e., physical violence, sexual violence, psychological violence, stalking, monitoring, and control) [2]. Each of these categories corresponds to specific violent behaviors which have been shown to change in their duration and severity[3]:

- Physical violence concerns the use of force to intentionally harm and injure the partner;
- Sexual violence refers to sexual acts or advances carried out without the victim's consent;
- Psychological violence corresponds to communication with the aim of detrimentally impacting the partner's mental and emotional well-being and exerting control over them;
- Stalking, monitoring, and control consists of persistent and unpleasant attention and communication inducing fear or concern about personal safety.

Moreover, the rising use of technologies has facilitated the escalation of the above described violent behaviors such that scholars have coined new forms of IPV ascribed to the so-called Cyber Intimate Partner Violence (C-IPV)[4]. C-IPV shares the same characteristics as IPV but occurs through the use of technologies or in cyberspace. Recurrent behaviors of C-IPV perpetrators include cyber sexual violence, cyber psychological violence, and cyber stalking, monitoring and control. Precisely, cyber sexual violence includes pressuring partners to send sexual content, coercing partners into sexual acts, and sending unwanted sexual content. Cyber psychological violence involves using technology, such as pictures, videos, and text messages, to cause emotional harm to partners, such as spreading rumours or insulting partners through text messages. Finally, cyber stalking, monitoring and control behaviors correspond to accessing electronic devices and accounts without permission to monitor their partner or have information on them. The majority of studies carried out so far provided useful information on the characteristics of these phenomena, their prevalence, individual differences (e.g., personality traits) correlated to the perpetration of both in-person and cyber IPV, and the detrimental consequences for victims [2, 5, 6]. In light of the detrimental consequences for victims of IPV and C-IPV, an imperative issue is trying to early detect these violent behaviors with the final goal of preventing their escalation. (C)-IPV detection can be problematic for victims because they are victims of their own romantic partner. In other words, being emotionally attached to the person who is committing violent acts towards themselves can reduce victims' ability to recognize such violent behaviors. Consequently, automatic detection of IPV and C-IPV behaviors can greatly help people in objectively identifying toxic and violent relationships and disengaging from them. This is the main motivation for our work: we propose the adoption of an LLM as an "assistant" being able to explain why a message, in the context of an intimate relationship, can be toxic. The explanation makes partners aware of the fact that violence is being committed or suffered and describes the reasons for this happening, as well as the consequences (for example, emotional suffering), with the hope that it can act as a deterrent.

## 3. LLaMAntino: an LLM for text generation in Italian Language

In this section, we briefly introduce the LLM used in our scenario. LLMs have proved their ability to excel in a large number of areas in the field of Natural Language Processing and also show good performance in solving tasks on which they have not explicitly been trained on [7, 8]. Notable examples of State-of-the-Art LLMs are surely represented by OpenAI's ChatGPT [9], Meta's LLaMA [10], BLOOM [11] and Mistral [12].

However, training these models requires an outstanding amount of computational resources and data for the training phases. This last requirement is particularly tricky in the case of languages other than English, which are known to be underrepresented. For the Italian language, there are other models in the literature, such as Camoscio [13] and Stambecco [14], both LLaMA instruction-tuned models, Fauno [15], a conversational Baize model and finally Cerbero [16], a Mistral-based model. All these models release few trained weights and do not exceed 13 billion in parameters.

LLaMAntino [17] is a family of LLMs that, starting from the pre-trained weights of LLaMA 2, were further refined for comprehension and text generation in the Italian language. The LLaMAntino training pipeline follows two main steps: the first one is represented by language adaptation, which allows a predominantly English model like LLaMA to adapt to the Italian language. The second step consists of fine-tuning the model to further improve its capabilities on specific tasks. Currently, the models

composing the LLaMAntino family are the following:

- **LLaMAntino-Chat** models based on the LLaMA 2-Chat versions[1] with language adaptation for Italian and further fine-tuning (7B, 13B, 70B).
- **LLaMAntino** models based on the LLaMA 2 versions[2] with language adaptation for Italian and instruction-tuning (7B, 13B, 70B).

Given these premises, we are now working on further fine-tuning LLaMAntino for downstream tasks like helping the user detect toxic behaviours and giving an explanation for its choice.

# 4. Explanations for Toxic Conversations

The idea is to create a dataset of toxic conversations annotated with information about the type of violence (e.g., physical, cyberstalking, cyber sexual violence), the presence of aggressive communication, the adoption of abusive language and, in general, with information that could be useful to provide a "technical" explanation, as if were given by a professional expert in the subject, such as a psychologist. The aim is to provide explanations, well grounded on relevant CIPV literature, that point out the elements of toxicity in the conversation. Therefore, we started from a dataset available on HuggingFace [18], which contains sentences classified as toxic or healthy, referring to teenage relationships. We extended the dataset by adding specific annotations related to CIPV to sentences classified as toxic. Then, we elaborated on the annotations to obtain an explanation that can be used for *Few-shot prompting*. The following subsections provide details on the dataset, annotation, and experiments.

## 4.1. Dataset and Annotations

The original dataset *"toxic-teenage-relationships"* was created to help in efforts to identify and curb instances of toxicity between teenagers[18]. It consists of 334 sentences collected by 8 teenagers (4 males and 4 females) of Spanish nationality aged between 15 and 19, who were appropriately instructed on interpersonal relationships to be classified as toxic or not. The group of teenagers had two weeks to collect Spanish language sentences that they spoke or heard in their environment either through interpersonal communication or via social media. Afterwards, the examples given by each student were discussed and evaluated by the others, using peer evaluation. The classification (toxic or non-toxic) was also approved by two specialists in the field. No personal or sensitive information has been recorded. As a general rule, if words associated with swearing, insults or profanity appear in a comment, it is likely to be classified as toxic, regardless of the author's tone or intention, e.g. humorous/self-critical. After classification, 165 sentences have been considered as toxic. With the aim of evaluating our Italian LLM, sentences have been translated into Italian by using two translation services (Google and DeepL). We added 5 of annotations:

- the type of violence: `physical` or `cyber`;
- the type of behavior that led to the physical violence, e.g. `sexual assault`, `stalking`;
- the type of cyber behavior that led to the violence, e.g. `cyber stalking`;
- the type of communication: `aggressive` or `non-aggressive`;
- the type of aggressive communication: e.g., `use of abusive language`.

As for physical violence, the experts distinguished 4 annotations [2]:

1. physical violence: the voluntary use of force that potentially causes harm and injury to the partner;
2. sexual violence: sexual acts without the partner's consent, even if only attempted;
3. psychological aggression: communicating with the intention of negatively influencing the mental and emotional state of the partner and wanting to control him or her;
4. stalking, monitoring and control: series of recurring and unwanted attentions and communications that create fear or apprehension and put the partner's safety at risk.

As for cyber violence, the experts distinguished 3 annotations [6]:

1. cyber sexual violence: requesting or pressuring the partner to send sexual content against his or her will, pressuring the partner to engage in sexual acts;
2. cyber psychological violence, aggression: behavior to cause emotional distress to the partner; may include behaviors such as spreading gossip on social media, repeatedly insulting the partner via messages, even spreading videos or photos that cause emotional distress;
3. cyber stalking, monitoring, and control: using and accessing technological devices and accounts without the partner's consent, use of technology to get information about your partner, in general any behaviours that aim at increasing control within the relationship). It includes *fraping*, that is the alteration of the partner's information on social profiles.

[1] https://huggingface.co/meta-llama/Llama-2-7b-chat
[2] https://huggingface.co/meta-llama/Llama-2-7b

As for aggressive communication, the experts distinguished 5 annotations [19]:

1. curses;
2. ridiculousness or derision;
3. bad language;
4. threat;
5. attack on the person (on competence, character, background, physical appearance).

At the end of the annotation phase, we had each toxic sentence annotated with information well-grounded in scientific literature about intimate partner violence. An example of a toxic sentence that reveals physical violence is:

> *"Tu non sei niente senza di me"* (*"You are nothing without me"*, in English)

That sentence has been annotated in the dataset as follows:

- type of violence: `physical`
- type of behaviour: `psychological aggression`
- aggressive communication: `yes`
- type of aggressive communication: `derision, attack on the person`

An example of a toxic sentence that reveals cyber violence is:

> *"Se non hai nulla da nascondere, dammi il telefono"* (*"If you have nothing to hide, give me your phone"*, in English)

which has been annotated in the dataset as follows:

- type of violence: `cyber`
- type of behaviour: `cyber stalking, monitoring, and control`
- aggressive communication: `yes`
- type of aggressive communication: `attack on the person`

The annotations will be exploited by LLM to generate explanations and raise awareness of the violent behaviour. In the next subsection, we describe how annotations are turned into examples for few-shot prompting.

## 4.2. Few-Shot Prompting to explain toxicity in conversations

We randomly chose 30 annotated toxic sentences for a small, preliminary experiment with Few-Shot Prompting; 20 sentences were used for training, 10 for testing. For each training sentence, the annotations were turned into a natural language explanation used to build prompts for in-context learning. For instance, the explanation for the previous sentence

> *"If you have nothing to hide, give me your phone"*

is: *"The sentence is toxic because it is an example of* `cyber violence`. *The behaviour falls in the category* `cyber stalking, monitoring, and control` *since the aim is to obtain information on the partner's life and establish a dynamic of control in the couple. Furthermore, the* `communication is aggressive` *because it reveals the intimidating intent of attacking the partner to violate his or her privacy."* We built a 2-shot prompt by including:

- the description of the task: "given a sentence from a conversation between partners in an intimate relationship, explain the reasons why the sentence expresses toxic language and represents a case of physical or cyber violence";
- 2 training toxic sentences with corresponding explanations;
- 1 test toxic sentence (without explanation) for which we want the model to generate an explanation.

In other words, the annotations associated with a toxic sentence were the canvas for writing the explanation included in the prompt. Therefore, we created 10 2-shot prompts, as described before, by using the 30 sentences extracted from the dataset. The aim of the experiment was to assess whether the annotations actually help in explaining the reasons why a message is classified as toxic. The model evaluated in our experiment was: **LLaMAntino-2-Chat-13B-hf-UltraChat**, LLaMAntino-2-Chat for brevity[3]. Therefore, we wanted to assess whether the model learns how to perform the task, by providing it with just two examples. We compared qualitatively the explanations given by LLaMAntino-2-Chat, when instructed by 2-shot prompts, with those generated when the model is prompted just with the task description and the toxic sentence to be explained ("zero-shot prompting"). The experimental protocol was:

1. give LLaMAntino-2-Chat the task description and the first toxic sentence to be explained and record the explanation;
2. repeat prompting with the remaining 9 test toxic sentences and record the explanations;
3. give LLaMAntino-2-Chat the 10 2-shot prompts and record the explanations;

After the generation step, for each test toxic sentence, we had 2 explanations: LLaMAntino-2-Chat 0-shot and LLaMAntino-2-Chat 2-shot. We asked 2 Psychology experts to evaluate independently the two explanations, by answering 3 questions:

[3] https://huggingface.co/swap-uniba/LLaMAntino-2-chat-13b-hf-UltraChat-ITA

**Table 1**
Answers given by experts on the 3 questions.

| | Expert 1 | | | Expert 2 | | |
|---|---|---|---|---|---|---|
| Answer | Q1 | Q2 | Q3 | Q1 | Q2 | Q3 |
| 0-SHOT | 40% | 40% | 60% | 60% | 50% | 70% |
| 2-SHOT | 40% | 50% | 30% | 40% | 20% | 30% |
| BOTH | 20% | 10% | 10% | 0% | 30% | 0% |
| NONE | 0% | 0% | 0% | 0% | 0% | 0% |

1. Q1: Which explanation is most scientifically based?
2. Q2: Which explanation is more effective in making the partner who suffers aware of the violence?
3. Q3: Which explanation is most effective for educational purposes to make both partners aware that violent behavior is taking place?

Explanations were presented in pairs. To avoid bias, experts are not aware of which training provided the explanation. Furthermore, the presentation order was random: sometimes the LLaMAntino-2-Chat 0-shot was presented before LLaMAntino-2-Chat 2-shot, sometimes the order was reversed. For each question, we suggested 4 possible outcomes: LLaMAntino-2-Chat 0-shot (anonymized), LLaMAntino-2-Chat 2-shot (anonymized), both, none. For each test sentence, we consider the experts to be in agreement if they gave the same answer to at least 2 of the 3 questions asked. In general, the expert were in agreement on 6 sentences, showing the difficulty of the task of evaluating the quality of explanations, given the sensitivity of the CIPV context.

Some interesting considerations have emerged from the results reported in Table 4.2, that can guide the next steps of the investigation:

- no question has ever been answered "none". Therefore, we can observe that the model never showed hallucinations or gave inappropriate answers. Of course, further testing will be necessary to generalize this statement;
- on Q1, the results suggest that there is no prompting strategy that clearly emerges, thus revealing that in general LLaMAntino-2-Chat explanations are properly based on scientific literature, regardless of the prompting strategy;
- on Q2, the answers show some disagreement among the experts: one was clearly in favour of LLaMAntino-2-Chat 0-shot, the other showed a slight preference for LLaMAntino-2-Chat 0-shot. We asked some motivations for the answers and it emerged that some explanations given by LLaMAntino-2-Chat 2-shot were negatively influenced by grammatical errors;
- on Q3, it seems that there is a clear evidence that LLaMAntino-2-Chat 0-shot explanations are more effective in making both partners aware of the violence.

In general, it seems that our LLM explains language toxicity with an adequate level of effectiveness, according to the 2 experts, but annotating sentences with information useful for few-shot prompting does not bring benefits on the explanations. This outcome might depend on the LLM used, as well as on the prompting strategy. Therefore, we plan to extend the experiment, obviously by increasing the size of the test set, comparing the results with another LLM, using Chain-of-Thought Prompting to improve the "reasoning" capabilities of the model.

## 5. Conclusions and Future Work

The prevalence of violent behaviors highlights the need for prompt intervention and preventive measures. We presented our proposal to utilize sophisticated Natural Language Processing techniques, including LLMs, to identify and describe toxic elements in discussions concerning teenage relationships. By exploiting the proficiency of LLMs in processing and understanding human language, our approach seeks to go beyond just the detection, aiming to grasp underlying motivations and factors contributing to the emergence of harmful behaviours.

In future works, we intend to perform fine-tuning steps to better adapt LLMs to the specific task at hand. We also plan to investigate how different pre-training techniques and architectures can be leveraged to enhance model performance. To ensure the effectiveness of our approach, we intend to confront our methodology with other models and incorporate further annotations to enhance the robustness and effectiveness of our methodology. This involves comparing the performance of our LLMs with other state-of-the-art models.

Moreover, we will explore the application of Chain-of-Thought prompting techniques, with the help of expert psychologists. This involves using prompts to guide the LLM's decision-making process, with the goal of encouraging the model to provide more detailed and grounded

explanations for its choices. By working closely with experts in this area, we hope to gain valuable insights into how these techniques can be best applied and refined. We plan also to extend the datasets with further annotations that provide more details about the language adopted (e.g. references to gender stereotypes or use of particular linguistic structures), with the aim of building more complete prompts.

## Acknowledgments

We acknowledge the support of the PNRR project FAIR - Future AI Research (PE00000013), Spoke 6 - Symbiotic AI (CUP H97G22000210007) under the NRRP MUR program funded by the NextGenerationEU. This Publication was produced with the co-funding of the European union - Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 - Partnerships extended to universities, research centres, companies and research D.D. MUR n. 341 del 15.03.2022 – Next Generation EU (PE0000014 - "SEcurity and Rights In the CyberSpace - SERICS" - CUP: H93C22000620001).

## References

[1] M. E. Bagwell-Gray, J. T. Messing, A. Baldwin-White, Intimate partner sexual violence: A review of terms, definitions, and prevalence, Trauma, Violence, and Abuse 16 (2015) 316–335.

[2] M. Breiding, K. C. Basile, S. G. Smith, M. C. Black, R. R. Mahendra, Intimate partner violence surveillance : uniform definitions and recommended data elements. version 2.0, 2015. URL: https://stacks.cdc.gov/view/cdc/31292.

[3] J. Spluska, L. Tanczer, Threat Modeling Intimate Partner Violence: Tech Abuse as a Cybersecurity Challenge in the Internet of Things, Emerald Publishing Limited, 2021, pp. 663–688.

[4] L. Gilbert, X. Zhang, K. Basile, M. Breiding, M.-j. Kresnow, Intimate partner violence and health conditions among u.s. adults —national intimate partner violence survey, 2010–2012, Journal of Interpersonal Violence 38 (2023) 237–261.

[5] K. N. Duerksen, E. M. Woodin, Cyber dating abuse victimization: Links with psychosocial functioning., Journal of Interpersonal Violence 36 (2021) NP10077–NP10105.

[6] L. Watkins, R. Benedicto, D. DiLillo, The cyber aggression in relationships scale: A new multidimensional measure of technology-based intimate partner aggression, Assessment 25 (2018) 608–626. doi:10.1177/1073191116665696.

[7] A. Tamkin, M. Brundage, J. Clark, D.-f. Ganguli, Understanding the capabilities, limitations, and societal impact of large language models, arXiv preprint arXiv:2102.02503 (2021).

[8] Y. Liu, T. Han, S. Ma, J. Zhang, Y. Yang, J. Tian, H. He, A. Li, M. He, Z. Liu, et al., Summary of chatgpt-related research and perspective towards the future of large language models, Meta-Radiology (2023) 100017.

[9] OpenAI, Gpt-4 technical report, 2023. arXiv:2303.08774.

[10] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, et al., Llama: Open and efficient foundation language models, arXiv preprint arXiv:2302.13971 (2023).

[11] B. Workshop, T. L. Scao, A. Fan, C. Akiki, E. Pavlick, S. Ilić, D. Hesslow, R. Castagné, A. S. Luccioni, F. Yvon, et al., Bloom: A 176b-parameter open-access multilingual language model, arXiv preprint arXiv:2211.05100 (2022).

[12] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. de Las Casas, F. Bressand, G. Lengyel, G. Lample, L. Saulnier, L. R. Lavaud, M. Lachaux, P. Stock, T. L. Scao, T. Lavril, T. Wang, T. Lacroix, W. E. Sayed, Mistral 7b, CoRR abs/2310.06825 (2023). URL: https://doi.org/10.48550/arXiv.2310.06825. doi:10.48550/ARXIV.2310.06825. arXiv:2310.06825.

[13] A. Santilli, E. Rodolà, Camoscio: an italian instruction-tuned llama, 2023. arXiv:2307.16456.

[14] Michael, Stambecco: Italian instruction-following llama model, https://github.com/mchl-labs/stambecco, 2023.

[15] A. Bacciu, G. Trappolini, A. Santilli, E. Rodolà, F. Silvestri, Fauno: The italian large language model that will leave you senza parole!, arXiv preprint arXiv:2306.14457 (2023).

[16] F. A. Galatolo, M. G. Cimino, Cerbero-7b: A leap forward in language-specific llms through enhanced chat corpus generation and evaluation, arXiv preprint arXiv:2311.15698 (2023).

[17] P. Basile, E. Musacchio, M. Polignano, L. Siciliani, G. Fiameni, G. Semeraro, Llamantino: Llama 2 models for effective text generation in italian language, arXiv preprint arXiv:2312.09993 (2023).

[18] Margarita Martínez Gabaldón, toxic-teenage-relationships (revision 5ce5df0), 2023. URL: https://huggingface.co/datasets/marmarg2/toxic-teenage-relationships. doi:10.57967/hf/0972.

[19] D. A. Infante, C. J. W. III, Verbal aggressiveness: An interpersonal model and measure, Communication Monographs 53 (1986) 61–69. doi:10.1080/03637758609376126.