Toward Discourse-Aware Models for Multilingual Fake News Detection

Francielle Vargas[1][2], Fabrício Benevenuto[2], Thiago A. S. Pardo[1]
[1]Interinstitutional Center for Computational Linguistics, University of São Paulo, Brazil
[2]Social Computing Laboratory, Federal University of Minas Gerais, Brazil
`francielleavargas@usp.br`
`fabricio@dcc.ufmg.br taspardo@icmc.usp.br`

Abstract

Statements that are intentionally misstated (or manipulated) are of considerable interest to researchers, government, security, and financial systems. According to deception literature, there are reliable cues for detecting deception and the belief that liars give off cues that may indicate their deception is near-universal. Therefore, given that deceiving actions require advanced cognitive development that honesty simply does not require, as well as people's cognitive mechanisms have promising guidance for deception detection, in this Ph.D. ongoing research, we propose to examine discourse structure patterns in multilingual deceptive news corpora using the Rhetorical Structure Theory framework. Considering that our work is the first to exploit multilingual discourse-aware strategies for fake news detection, the research community currently lacks multilingual deceptive annotated corpora. Accordingly, this paper describes the current progress in this thesis, including (i) the construction of the first multilingual deceptive corpus, which was annotated by specialists according to the Rhetorical Structure Theory framework, and (ii) the introduction of two new proposed rhetorical relations: INTERJECTION and IMPERATIVE, which we assume to be relevant for the fake news detection task.

1 Introduction

According to DePaulo et al. (2003), liars may be identified by their words, and a fairly straightforward element to mitigate risks of deceptive activities consists of the ability to correctly interpret and codify the underlying intent of the speakers (Ho and Hancock, 2019).

Several models have proposed to automatically identify statements that are intentionally misstated (or manipulated). The vast majority of deception detection methods rely on linguistic features such as n-grams, language complexity, part-of-speech tags, and syntactic and semantic features.

More recently, discourse frameworks of the computational linguistics area, such as Rhetorical Structure Theory (RST) (Mann and Thompson, 1987), have been adopted as the basis to automatically detect deceptive stories (e.g., fake news or fake reviews) on the web and social media. Figure 1 shows an example of a fake news annotated using the Rhetorical Structure Theory framework.
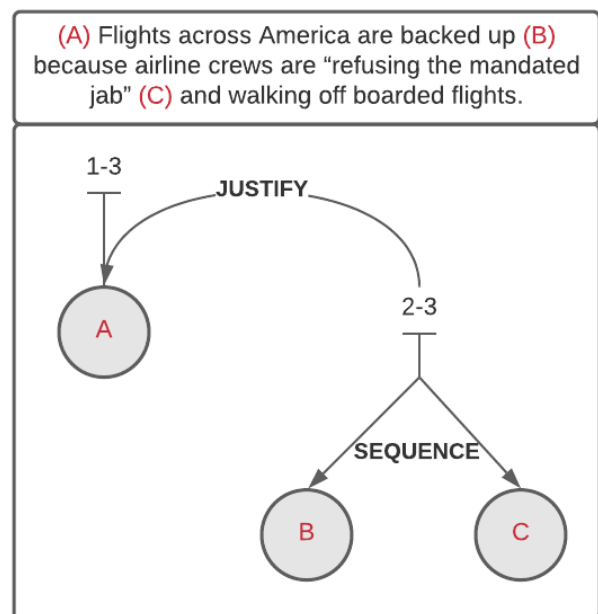

Figure 1: Example of a fake news extracted from Politifact stated on August 8, 2021 in a Facebook post, which was annotated using Rhetorical Structure Theory framework.

As shown in Figure 1, a deceptive story (also called fake news or deceptive news) was segmented into three elementary discourse units (EDU's), as well as hierarchically organized according to the following coherence relations: JUSTIFY and SEQUENCE. Moreover, there are spans of texts that

are more central (nucleus) to the text's purpose than others (satellite), and the nucleus is signaled by an arrow.

Kumar and Geethakumari (2014) argue that research on people's cognitive mechanisms has promising guidance for the detection and refutation of fake content. Nevertheless, most approaches for fake news detection are focused on the content-based classification, as well as "shallow" linguistic features, even without considering the potential of cognitive mechanisms or contextual information. Therefore, taking into account the considerable impact of discourse and contextual information for deception detection, we allow ourselves to argue that would be a carelessness rely not on rich and more sophisticated approaches for the fake news detection task.

Although there is an acknowledged potential for a discourse-aware analysis as a cognitive-based approach, there is a considerable lack of research on discourse in deceptive stories. To Meibauer (2018), we do not have significant knowledge of the embedded lies in text or discourse, with the notable exception of the studies proposed in Galasińki (2000) and Meibauer and Dynel (2016), although these studies deal with fictional discourse. Indeed, a very plausible assumption to explain the lack of research on deception in real discourses would be that this is rather a difficult requirement to satisfy.

More recent proposals have explored discourse-aware models for deception detection, more specifically, fake news and fake reviews detection, nevertheless, to the best of our knowledge, no previous multilingual discourse-aware studies exist. The first discourse approach for fake news detection was proposed by Rubin et al. (2015), in which a discourse-aware model that embodies rhetorical structure features using classical learning methods was used for the English language. In the same settings, for the Russian language, Kuzmin et al. (2020), and Pisarevskaya (2017) also developed a classical Bag-of-RST representation using Support Vector Machine (SVM) (Scholkopf and Smola, 2001) and Random Forest (Breiman, 2001), as well as pre-trained contextual embeddings (e.g., BERT) were used for fake news detection. In Atanasova et al. (2019), a robust framework was proposed, which embodies discourse features extracted from a RST-discourse parser, using Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) and Convolutional Neural Networks (CNN)

sophisticated learning methods for fake news detection in English. The authors claim that discourse and contextual features improve the fake news detection system performance. For fake reviews detection, Popoola (2017) analyzes rhetorical structure coherence relations from a forensic collection of authentic and fake Amazon book reviews and concludes that paid review writers deploy deceptive pragmatics (i.e., a coherent set of linguistic strategies) are deployed to support the intent to deceive.

In this thesis, we embrace the challenges and opportunities of the study on the rhetorical structure patterns in fake news from different languages. The main goal of this thesis is to assess whether, from the study on the discourse structure patterns in a multilingual fake news corpora, it is possible to propose efficient computational discourse-aware models for multilingual fake news detection. Accordingly, this investigation is an inquiry into the nature of the high-level, discourse structure of deceptive natural language texts, whereby the long-term main goal is to uncover systematic language differences and inform them for deception or falsification verification systems.

In what follows, we will give a summary of our research effort so far. In Section 2, we briefly discuss decetive language, as well the state-of-the-art on discourse structure applied to deception detection. Moving forward, in Section 3, we present this ongoing research and describe the progress toward this thesis's goals by (1) the construction of the first multilingual deceptive corpus annotated according to the Rhetorical Structure Theory framework, and (2) introducing two new proposed rhetorical relations: INTERJECTION and IMPERATIVE, which we suggest to be relevant for fake news detection. Section 4 describes the future research paths for further exploration. In Section 5, some conclusions are presented.

2 Deceptive Language and Discourse

Information is mainly passed on in the form of language. Consequently, research on how people use language to inform and misinform, as well as how meaning may be automatically extracted from large amounts of text, is unquestionably relevant. Moreover, the main literature on deception suggests that liars may be identified by their stories (Newman, 2003; DePaulo et al., 2003; Hancock and Guillory, 2015; Ho and Hancock, 2019).

According to DePaulo et al. (2003), deceptive linguistic style in general presents negative polarity, weak employment of singular and third-person pronouns, and high employment of movement verbs. In the same settings, Nahari et al. (2019) claim that the basic assumption on deceptive language is that liars differ from truth-tellers in their verbal behavior, accordingly, lying stories may be identified by inspecting the verbal behavior patterns. In the same setting, Newman (2003) proposes a set of linguistic behaviors, which, according to the authors, predict deception, such as tones of words, kind of preposition, conjunctions, and pronouns.

A very plausible assumption would be that there are different discourse structure patterns in deceptive and truthful stories. Indeed, this has been claimed by various authors (Rubin et al., 2015; Popoola, 2017; Kuzmin et al., 2020; Pisarevskaya, 2017; Atanasova et al., 2019). Discourse-aware approaches for deception detection are usually framed as a supervised learning problem, which embodies in a model coherence relations followed by hierarchical nuclearity information to build automatic classifies. Moreover, the discourse features may be also used as a complement for other stylistic-based linguistic features. Therefore, discourse analysis applied to automated deception detection intends to assess the potential for the application of discourse coherence analysis, developing linguistic heuristics to automatically classify statements that are intentionally misstated (or manipulated).

3 Ongoing Research

In this Ph.D. ongoing research, we are especially interested in discourse-aware approaches for deception detection, more specifically, using the Rhetorical Structure Theory framework for fake news detection. Our aim is to develop computational models for deception detection where discourse structure is a first-class citizen. We have already made progress toward this goal by: (1) developing the first multilingual deceptive corpus, which we called Deceiver [1], annotated by specialists according to the RST framework. Deceiver multilingual corpus consists of 600 pieces of news classified in deceptive versus truthful documents, as well split in 300 documents in Brazilian Portuguese, and 300 documents in English; and (2) introducing two new rhetorical relations: INTERJECTION and IMPER-

ATIVE, which we consider to be relevant for the fake news detection task.

Moving forward, we propose to examine the discourse structure. Our aim is to develop patterns that show deception. Secondly, we will assess whether a computational model that embodies the multilingual discourse structure patterns is an efficient approach for multilingual fake news detection on the web and social media. Therefore, the proposed research is divided among three main fronts described in what follows and shown in Figure 2:

- Multilingual Corpus Construction;

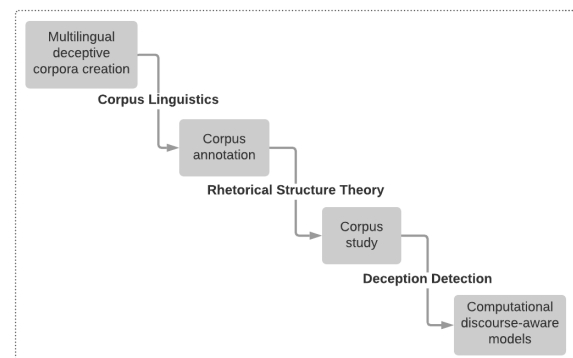- Rhetorical Structure Theory Modeling;

- Deception Detection.

Figure 2: Methodology

3.1 Multilingual Corpus Construction

As our work is the first to explore multilingual discourse-aware strategies for deception detection, the research community currently lacks multilingual deceptive annotated corpora. Therefore, we provide a new corpus composed of fake news and truthful stories in Portuguese and English, which was annotated by two skilled linguistics using the Rhetorical Structure Theory framework.

3.1.1 Corpus Development

According to Bachenko et al. (2008), obtaining reliable positive and negative data samples is one of the challenges in automated deception detection research, especially concerning deception detection tasks that require careful selection of data. The difficulty is in-ground truth verification: finding suitable data "in the wild" and conducting fact checking to obtain ground truth is costly, time-consuming, and labor-intensive. Furthermore, as

---

[1]https://github.com/francielleavargas/Deceiver

a matter of ongoing debate, exactly what counts as a representative corpus is an open question and difficult to satisfy.

In this thesis, we created and annotated the first multilingual deceptive corpus, which we named "Deceiver" [2], which consists of deceptive and truthful news shared on the web and social media. The Deceiver multilingual corpus consists of 600 pieces of truthful and deceptive news divided into 300 pieces of news in Brazilian Portuguese, and 300 pieces of news in English classified equally into deceptive and truthful classes. Table 1 exhibits the corpus overview.

Table 1: Deceiver multilingual corpus overview

Deceiver Corpus	Data	Total	Language
Fake News	New stories	150	Portuguese
Truthful News	CSTnews	150	Portuguese
Fake News	MisInfoText	150	English
Truthful News	MisInfoText	150	English

New stories consists of unpublished deceptive stories manually collected from 3 (tree) Brazilian fact-checking agencies: G1 Fato ou Fake[3], Lupa[4] and E-farsas[5]. The collected stories were checked by the mentioned agencies during the period of July-2018 to June-2021. We also used a couple of pieces from another fake news corpus in Brazilian Portuguese proposed by our research laboratories.

CSTnews (Cardoso et al., 2011) consists of a corpus composed of 140 news texts written in Brazilian Portuguese and counts with several annotation layers, including annotation in the morphosyntax/syntax, semantics, and discourse levels. The annotation has been performed by computer scientists and linguists. We revisited the discourse annotation level of the CSTnews corpus and included the annotation of new coherence relations proposed by this research (see Sections 3.2.1 and 3.2.2). Note that CSTnews provides only 140 news texts. Therefore, we additionally collected 10 pieces of truthful news written in the Brazilian Portuguese language from Folha de São Paulo [6], which consists of a referential Brazilian newspaper. The new documents were randomly collected during July 2021 on issues related to coronavirus pandemic topics.

MisInfoText (Torabi Asr and Taboada, 2018) consists of a large dataset containing instances of fact-checked news articles, which were automatically collected from fact-checking websites for false/true claims and headlines, links to the original news articles spreading them, and the veracity labels given by the fact-checkers. We randomly selected 150 fake and 150 real stories and annotated them using rhetorical structure theory.

3.1.2 Corpus Annotation

Consistency and quality of the data are directly related to the performance of the machine learning classifiers. In particular, subjective tasks, such as fake news and fake reviews detection present high complexity and a wide range of challenges, and an annotation schema provides a suitable characterization of domain specificity, as well as improves the labeled data.

In this thesis, we annotated 600 deceptive and truthful stories using the Rhetorical Structure Theory framework. The annotation process was performed by two specialists and conducted by an experienced linguist, in order to warrant the quality of the labeled data, as well as to minimize the subjectivity of this task, including the natural ambiguity related to RST coherence relations classification. Accordingly, we propose an annotation approach, which encompasses three main steps: (i) selection of annotators, (ii) annotation schema definition, and (iii) evaluation, which we describe in what follows:

1. **Selection of annotators**: firstly, we selected annotators with suitable background. Given that the RST is a linguistic theory, we selected two skilled linguists with backgrounds in the RST framework. Table 2, exhibit the specialists profile.

Table 2: Specialist's profile.

Degree	Occupation	Knowledge in Portuguese	Knowledge in English
Linguistics	Professor of Syntax	Mother language	Proficient
Linguistics	NLP Specialist	Mother language	Proficient

2. **Annotation schema definition**: we propose a well-defined annotation schema to minimize mistakes arising from human subjectivity, as well as leverage the annotation precision. We use the Rhetorical Structure Theory framework, and supply a well-written and clear guideline in the specialist's mother language,

which guided the annotators during the entire annotation process. Our guideline was build using the RST-annotation guideline proposed by Stede et al. (Stede et al.) and Carlson and Marcu (2001), as well as the original RST proposal Mann and Thompson (1987); Thompson and Mann (1988). In total, 38 (thirty-eight) rhetorical relations were annotated, being 32 (thirty-two) rhetorical relations proposed in Mann and Thompson (1987), 4 (four) rhetorical relations proposed in Carlson and Marcu (2001), and 2 (two) new rhetorical relations identified during this ongoing thesis research. Table 3 describes the set of rhetorical relations used to annotate the proposed Deceiver multilingual corpus.

Table 3: Rhetorical relations used to annotate the Deceiver multilingual corpus

Author	Total	Description
(Mann and Thompson, 1987)	32	ANTITHESIS, BACKGROUND, CIRCUMSTANCE, CONCESSION, CONDITION, ELABORATION, ENABLEMENT, EVALUATION, EVIDENCE, INTERPRETATION, JUSTIFY, MOTIVATION, NON-VOLITIONAL CAUSE, NON-VOLITIONAL RESULT, OTHERWISE, PURPOSE, RESTATEMENT, SOLUTIONHOOD, SUMMARY, VOLITIONAL CAUSE, VOLITIONAL RESULT, CONTRAST, SEQUENCE, JOINT, ATTRIBUTION, COMPARISON, PRESENTATION, DISJUNCTION, MEANS, PREPARATION, UNCONDITIONAL, UNLESS.
(Carlson and Marcu, 2001)	4	LIST, CONCLUSION, PARENTHETICAL, SAME-UNIT
Proposed Relations	2	INTERJECTION, IMPERATIVE

Moreover, we selected easy-accessible platforms for discussion among annotators (e.g., WhatsApp and e-mail), where the annotators could discuss aspects related to the chosen theory, as well as on the ambiguities inherent in natural languages.

3. **Evaluation**: Our evaluation approach is divided into two steps. In the first step, 45% of the corpus was annotated by specialist 1, and other 45% of the corpus was annotated by specialist 2. In the second step, 10% of the corpus was annotated by both annotators. We evaluated the annotation process using this sample of the full corpus, composed of 10% of documents, which were annotated by both linguists. Figure 3 shows our evaluation approach.
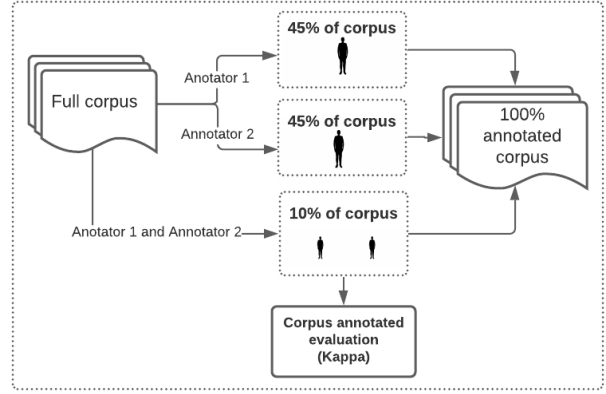

Figure 3: Evaluation

Currently, the annotation evaluation is ongoing. We further intend to apply the Kappa (McHugh, 2012) to measure inter-human annotation agreement.

3.2 Rhetorical Structure Theory

Mann and Thompson (1987) claim that Rhetorical Structure Theory is "a linguistic theory of how text hangs together, and extended discourses are made out of sequences of sentences that are related to one another". RST presents three mechanisms that are central above all for understanding the proposed framework. Mann and Thompson (1987) define these mechanisms by nuclearity, schema, and rhetorical (or coherence) relations.

Rhetorical relations contain nuclei and satellites. The nucleus consists of text elements that are more central and relevant to the relation. On the other hand, the supporting units are called satellites. Coherence relations are also classified in mononuclear and multinuclear relations, and described in terms of the used schema (i.e., how one or more satellites or nuclei relate to each other).

In this paper, we present two unpublished rhetorical relations, which were identified during the annotation process of the first multilingual deceptive corpus. In Sections 3.2.1 and 3.2.2, we describe the 1INTERJECTION and IMPERATIVE rhetorical relations, that we suggest to be relevant to signal deceptive intentions.

3.2.1 INTERJECTION

We agree that natural languages are ultimately socially transmitted. Accordingly, Dingemanse (2017) claims that the structure of language forms the evolutionary landscape for linguistic items, and there are unrefined linguistic structures that signal characteristics of the margins of language. Wilkins

(1992) defines this type of linguistic structure as conventional lexical forms that are monomorphemic and typically constitute an utterance of their own. Therefore, in natural languages, an interjection also may be defined as an expression that transmits sudden emotions.

As already mentioned, deceptive news may be analogous to satires, as well as emotions are noted. Moreover, despite that RST provides some coherence relations that supply pragmatic information, no previous coherence relations exist to signal intentions that manifest emotions, sensations, and states of mind using unrefined linguistic structures. Therefore, in this thesis, we assume that a new rhetorical relation provides relevant information for deception detection tasks, and propose a new mononuclear relation named INTERJECTION, described in Table 4 in terms of constraints on the nucleus, satellite, and effects on the reader.

Table 4: INTERJECTION rhetorical relation: brief description of the constraints on the nucleus and satellites, nuclearity classification, and type of relation (information or intentional). We should point out that *N/A* means that there are not any constraints.

Type	Description
Brief description	INTERJECTION is a type of rhetorical relation that signals the intent to act on the reader, leading him to adopt a type of behavior, making use of the unrefined linguistic structures that express information analogous to sudden emotions, sensations, and states of mind.
Type of nucleus	Mononuclear.
Type of relation	Intentional.
Constraints on the nucleus (N)	N/A
Constraints on the satellite (S)	The satellite presents any context that provides a piece of information with sudden emotions, sensations, states of mind.
Constraints on the combination between nucleus and satellite (N+S)	N/A
Effect on the reader	Writer's intention consists of leading the reader to adopt a certain behavior without the need to make use of more elaborate linguistic structures.

As shown in Table 4, the proposed INTERJECTION relation is mononuclear and presents pragmatic information. The satellite presents necessarily uncomplicated linguistic structures that express information analogous to sudden emotions, sensations, and states of mind information. Moreover, there are not any constraints on the nucleus, and the

writer leads the reader to adopt a type of behavior using an unrefined linguistic structure analogous to sudden emotion, sensations, and states of mind. An example of INTERJECTION is shown in Figure 4.
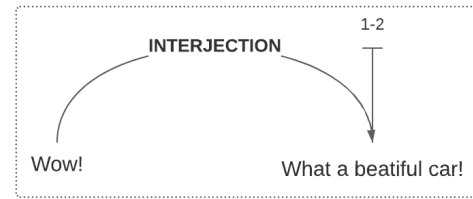

Figure 4: INTERJECTION rhetorical relation.

3.2.2 IMPERATIVE

Levels of analysis followed by different linguistic structures encompass the information passed in natural languages. Linguists around the world have been spending thousands of centuries understanding and proposing theories that explain the complex human language phenomena. An elementary fundamental of language structure is that the vocabulary of most languages is classified according to word-class or grammatical categories. The languages usually present the following distinctive word classes: noun, adjective, pronoun, verb, article, numeral, adverb, conjunction, and interjection. Each one of these categories is responsible for signaling events into natural language. For instance, note that nouns are responsible for naming the objects and beings in the world. Differently, the actions performed by the objects and beings in the world are usually signaled by verbs. Therefore, verbs are an important word class that expresses an action, state, or process. Furthermore, a wide variety of verbs may supply discourse and pragmatics information. For instance, observe the following sentence:

1. *I insist that he leaves the house right now.*

2. *Get out of the house right now!*

Note that, in the second example, we observe that the writer intends explicitly to give an order in an authoritative way. On the other hand, in the first example, even though there is a typical order intentionally signaled by the verb "insist", it is not expressed explicitly in an authoritative way. Taking into account the RST framework, the first sentence would be classified as a MOTIVATION rhetorical relation. Moving forward, since this type of coherence relation was identified in the deceptive

stories, we propose a new rhetorical relation named IMPERATIVE. Table 5 presents the definition that we propose to this relation.

Table 5: IMPERATIVE rhetorical relation: brief description of the constraints on the nucleus and satellites, nuclearity classification, and type of relation (information or intentional). We should point out that "N/A" means that there are not any constraints.

Type	Description
Brief description	IMPERATIVE rhetorical relation expresses order, request or advice, whose writer imposingly intends that the reader accomplish an action. Moreover, this type of rhetorical relation may be found in an affirmative or negative way.
Type of nucleus	Mononuclear.
Type of relation	Intentional.
Constraints on the nucleus (N)	N/A.
Constraints on the satellite (S)	The satellite expresses an explicit order, advice, or request.
Constraints on the combination between nucleus and satellite (N+S)	The nucleus precedes the satellite in the text.
Effect on the reader	The writer intends explicitly to give an order in an authoritative way. The reader recognizes the presence of an order, advice, or request.

As shown in Table 5, the proposed IMPERATIVE relation is mononuclear and presents pragmatic information. The satellite expresses an explicit order, advice, or request, and there are not any constraints on the nucleus. On the effects on the reader, the reader understands that there is the presence of order, advice, or request. Figure 5 exhibits an example of IMPERATIVE rhetorical relation.
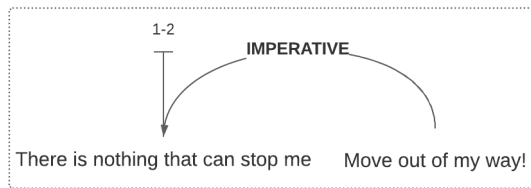

Figure 5: IMPERATIVE rhetorical relation.

4 Further Research Exploration

4.1 Deception Detection

In the last few years, there was a growth in the number of web and social media users. Consequently, the potential of deceptive activities also increased, resulting in the production of fake news, fake reviews (also known as opinion spam), deceptive

discussion, simple lies, etc. Fake news detection is defined as the prediction of the chances of a particular news article being intentionally deceptive (Rubin et al., 2015). Fake reviews or opinion spams are inappropriate or fraudulent reviews (Ott et al., 2011). Deceptive discussions consist of narratives or statements intentionally misstated (or manipulated) (Larcker and Zakolyukina, 2012).

As it is known from previous discussions, Rhetorical Structure Theory has been adopted for fake news detection. Accordingly, in this Ph.D. ongoing research, we propose to examine rhetorical structure patterns from the proposed multilingual deceptive corpus and to build and evaluate computational discourse-aware models for multilingual fake news detection on the web and social media. Moving forward, the next steps of this thesis include (a) performing a corpus study with the proposed RST-annotated multilingual deceptive corpus, and (b) building computational models, which embody discourse structure information, and evaluating whether our approach is efficient to automatically detect fake news in different languages. In Sections 4.1.1 and 4.1.2, we discuss the future research paths for further exploration in this thesis.

4.1.1 Corpus Study

The next step of this ongoing research consists of performing a skilled corpus study, in which a specialist will theoretically and empirically analyze linguistic patterns in deceptive and truthful news, in order to identify linguistic structures that show deception. Firstly, we will analyze lexical, syntactic, and semantic phenomena in deceptive and truthful stories, and assess the relationship between these phenomena in both analyzed languages: Portuguese and English.

Secondly, we propose to examine the discourse structure in deceptive and truthful stories in order to assess whether deceptive and truthful discourse structure presents similar multilingual patterns. In order to achieve that, we intend to investigate different types of coherence relations, as well as the nuclearity properties.

4.1.2 Building Discourse-Aware Computational Models

Recently proposed approaches in the literature assess discourse structure from deceptive and truthful stories, and build models, which embody discourse structure information for monolingual deception detection. Differently from previous discourse-based

approaches for fake news detection, we intend to examine discourse structures from deceptive and truthful stories for multilingual fake news detection. In order to archive that, the next step of this ongoing research consists of building computational discourse-aware models, which embody discourse structure patterns learned from a multilingual deceptive corpus study, as well as evaluating whether the proposed approach may efficiently distinguish fake news from real news in different languages.

In general, previous work on fake news detection uses neural networks or traditional machine learning techniques for specific languages (Rubin et al., 2015; Kuzmin et al., 2020; Pisarevskaya, 2017; Atanasova et al., 2019). In particular, due to the high performance presented by the deep artificial neural networks, which differs from traditional machine learning techniques due to the learning of abstract feature representations through its multiple layers, we intend to evaluate classical machine learning classifiers and employ deep neural network learning methods and pre-trained models of word embeddings, such as BERT (Devlin et al., 2019).

5 Conclusions

Statements that are intentionally misstated (or manipulated) present linguistic features that are different in relation to the statements that are not intentionally misstated. According to Fallis (2009), lying is saying something that you believe to be false with the intent to deceive, as well as deceptive language consists of a type of language deliberately used to attempt to mislead others (DePaulo et al., 2003). In recent years, there is a considerable increase in deceptive activities on the web and social media, including, for example, fake news, which consist in a global and relevant social, political, and economic problem.

Prior researches on fake news detection use a variety of linguistic styles and psychologist-based features for monolingual and multilingual fact-checking systems. Rather than using the classical linguistic style approach, recent research works propose successfully discourse-aware computational models for predicting online deception, more specifically, fake news and fake reviews. Nevertheless, to the best of our knowledge, there are not any multilingual discourse-aware studies on deceptive texts. Therefore, in this Ph.D. ongoing research, we embrace the challenges and opportunities on the

study of multilingual discourse structure in deceptive natural language texts for prediction of online deception, more specifically, fake news.

We aim further to investigate discourse structure patterns from the first proposed multilingual deceptive corpus, annotated according to the Rhetorical Structure Theory framework. Then, we intend to build computational models that embody the discourse structure patterns that show deception, as well as to evaluate whether our approach is efficient for the multilingual fake news detection task.

In this paper, we presented the ongoing Ph.D. research and described the already in progress work toward this thesis's goals. Ours research method includes (a) the construction of the first multilingual corpus of deceptive versus truthful stories (news) in Portuguese and English languages, and (b) the introduction of two new rhetorical relations, which consist of the second contribution of this research. The proposed rhetorical relations, named INTER-JECTION and IMPERATIVE, present pragmatic information and are mononuclear. Moreover, we suggest that these proposed relations are relevant for discourse-aware approaches for fake news detection. Finally, we discuss the future research paths for further exploration related to this thesis's goals.

References

Pepa Atanasova, Preslav Nakov, Lluís Màrquez, Alberto Barrón-Cedeño, Georgi Karadzhov, Tsvetomila Mihaylova, Mitra Mohtarami, and James Glass. 2019. Automatic fact-checking using context and discourse information. *J. Data and Information Quality*, 11(3).

Joan Bachenko, Eileen Fitzpatrick, and Michael Schonwetter. 2008. Verification and implementation of language-based deception indicators in civil and criminal narratives. In *Proceedings of the 22th International Conference on Computational Linguistics*, pages 41–48, Manchester, UK.

Leo Breiman. 2001. Random forests. *Machine Learning*, 45:5–32.

P.C.F. Cardoso, E.G. Maziero, M.L.R. Castro Jorge, E.M.R. Seno, A. Di Felippo, L.H.M. Rino, M.G.V. Nunes, and T.A.S. Pardo. 2011. Cstnews - a discourse-annotated corpus for single and multi-document summarization of news texts in brazilian portuguese. In *Proceedings of the 3rd RST Brazilian Meeting*, pages 88–105, Cuiabá, Brazil.

Lynn Carlson and Daniel Marcu. 2001. Discourse tagging manual. *Tech. rep. ISI-TR-545*, pages 01–87.

Bella DePaulo, James J Lindsay, Brian Malone, Laura Muhlenbruck, Kelly Charlton, and Harris Cooper. 2003. Cues to deception. *Psychological Bulletin*, 129:74–118.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.*, pages 4171–4186, Minnesota, USA.

Mark Dingemanse. 2017. On the margins of language: Ideophones, interjections and dependencies in linguistic theory. In N. J. Enfield, editor, *Dependencies in language*, volume 15, page 195–203. Language Science Press.

D. Fallis. 2009. What is lying? *Journal of Philosophy*, 106(1):29–56.

Dariusz Galasiński. 2000. *The language of deception: a discourse analytical study*. SAGE Publications.

Jeffrey T. Hancock and Jamie Guillory. 2015. Deception with technology. In *The Handbook of the Psychology of Communication Technology*, volume 16, pages 270–289.

Shuyuan Mary Ho and Jeffrey T. Hancock. 2019. Context in a bottle: Language-action cues in spontaneous computer-mediated deception. *Computers in Human Behavior*, 91:33–41.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.

KP Krishna Kumar and G Geethakumari. 2014. Detecting misinformation in online social networks using cognitive psychology. *Human-centric Computing and Information Sciences*, 4(14).

Gleb Kuzmin, Daniil Larionov, Dina Pisarevskaya, and Ivan Smirnov. 2020. Fake news detection for the Russian language. In *Proceedings of the 3rd International Workshop on Rumours and Deception in Social Media*, pages 45–57, Barcelona, Spain.

David F. Larcker and Anastasia A. Zakolyukina. 2012. Detecting deceptive discussions in conference calls. *Journal of Accounting Research*, 50(2):495–540.

William C. Mann and Sandra A. Thompson. 1987. *Rhetorical structure theory: a theory of text organization*. University of Southern California, Information Sciences Institute Los Angeles.

Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.

Jörg Meibauer. 2018. The linguistics of lying. *Annual Review of Linguistics*, 4(1):357–375.

Jörg Meibauer and Marta Dynel. 2016. Empirical approaches to lying and deception. *International Review of Pragmatics*, 8(3).

Galit Nahari, Tzachi Ashkenazi, Ronald Fisher, Pär Granhag, Irit Hershkowitz, Jaume Masip, Ewout Meijer, Zvi Nisin, Nadav Sarid, Paul Taylor, Bruno Verschuere, and Aldert Vrij. 2019. Language of lies: Urgent issues and prospects in verbal lie detection research. *Legal and Criminological Psychology*, 24:1–23.

Pennebaker J. W. Berry D. S. Richards J. M. Newman, M. L. 2003. Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 5(29):665–675.

Myle Ott, Yejin Choi, Claire Cardie, and Jeffrey T. Hancock. 2011. Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 309–319, Oregon, USA.

D Pisarevskaya. 2017. Rhetorical structure theory as a feature for deception detection in news reports in the russian language. In *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference Dialogue 2017*. ACM.

Olu Popoola. 2017. Using rhetorical structure theory for detection of fake online reviews. In *Proceedings of the 6th Workshop Recent Advances in RST and Related Formalisms*, pages 58–63, Santiago de Compostela, Spain.

Victoria L Rubin, Niall J Conroy, and Yimin Chen. 2015. Towards news verification: Deception detection methods for news discourse. In *Proceedings of the The Rapid Screening Technologies, Deception Detection and Credibility Assessment Symposium co-located with HICSS48*, page 01–11, Hawaii, USA.

Bernhard Scholkopf and Alexander J Smola. 2001. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, Cambridge.

Manfred Stede, Maite Taboada, and Debopam Das. 2017. Annotation guidelines for rhetorical structure. *Linguistics Department at The University of Potsdam*, 1:1–31.

Sandra A. Thompson and William C. Mann. 1988. Rhetorical structure theory: A framework for the analysis of texts. *IPRA Papers in Pragmatics*, 1:79–105.

Fatemeh Torabi Asr and Maite Taboada. 2018. The data challenge in misinformation detection: Source reputation vs. content veracity. In *Proceedings of the First Workshop on Fact Extraction and VERification (FEVER)*, pages 10–15, Brussels, Belgium.

David P. Wilkins. 1992. Interjections as deictics. *Journal of Pragmatics*, 18(2-3):119–158.