Automatic Detection of Contradictions in Texts

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
der Philosophie des Fachbereiches 05 – Sprache, Literatur, Kultur
der Justus-Liebig-Universität Gießen

vorgelegt von
Natali Karlova-Bourbonus, M.A.

Aus Frankfurt am Main

2018

Vorsitzender: Herr Prof. Dr. Thomas Möbius

Erstgutachter: Herr Prof. Dr. Henning Lobin

Zweitgutachter: Herr Prof. Dr. Helmut Feilke

For my little prince Philipp

Acknowledgments

Five years are now passed since I began my doctoral thesis on contradictions in news texts and at this point, I would like to thank everybody who helped me in finalizing it.

First of all, I would like to express my special thanks of gratitude to my *Doktorvater*, Prof. Dr. Henning Lobin, who gave me a priceless opportunity to collect new experience in teaching and researching and to learn how exciting natural language processing can be. Dear Henning, thank you for supporting me the whole, often thorny, way to the completed monograph from the first half-baked ideas till the last dot on the paper.

I would also like to express my special thanks of gratitude to my second supervisor, Prof. Dr. Helmut Feilke. Participation in his *GCSC-Kolloquium* in the year 2014 led to a long-desired breakthrough in my research. I am grateful for his sharing of precious ideas with me, which laid the groundwork for my research.

While reading the present monograph you will come across the statement several times that finding contradictions is a challenging task for a human. I would like to thank all the students at the University of Giessen who did not let this challenge discourage them and who, in the years 2015 and 2016, took part on my surveys conducted within the framework of the present study. Their patience, curiosity, and ability to explore latent things in the texts are admirable.

Last but not least, I would like to thank the most important people in my life – my family. My parents, who always let me go my way, who never judged me for my decisions, and patiently and gratuitous supported me in seemingly inextricable moments. My husband Nico, who inspired and motivated me to make possible what, in the beginning, seemed to be impossible. And my brother Denis for the exciting conversations on the processing of natural languages.

Table of Contents

List of Abbreviations

ACE	Automatic Content Extraction Programm
ACL	Association for Computational Linguistics
BART	Baltimore Anaphora Resolution Toolkit, currently Beautiful Anaphora Resolution Toolkit
CBOW	Continuous Bag of Words
CCG	Componential Counting Grid
CD	Contradiction Detection
CFG	Context-free Grammar
COLING	International Conference on Computational Linguistics
CoNLL	The SIGNLL Conference on Computational Natural Language Learning
CoU	Context of Utterance
DIRT	Discovery of Inference Rules from Text
DNA	Duplex Negatio Affirmat
DNN	Duplex Negatio Negat
DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
DRT	Discourse Representation Theory
DSM	Distributional Semantic Model
EDITS	Edit Distance Textual Entailment Suite
EMD	Earth Mover's Distance
EOP	EXCITEMENT Open Platform
ESA	Explicit Semantic Analysis
EXCITEMENT	EXploring Customer Interactions through Textual EntailMENT
FOL	First-Order Logic
FraCas	A Framework for Computational Semantics
GLSA	Generalized Latent Semantic Analysis
HAL	Hyperspace Analogue to Language
HLBL	Hierarchical Log-bilinear Model
HOTCoref	Higher Order Tree Coreference

kNN	k-Nearest-Neighbor
LDA	Latent Dirichlet Allocation
LDC	Linguistic Data Consortium
LEM	Law of Excluded Middle
LNC	Law of Non-Contradiction
LSA	Latent Semantic Analysis
LSI	Latent Semantic Indexing
mSDA	Marginaized Stacked Denoising Autoencoders
MUC	Message Understanding Conference
NC	Negative Concord
NER	Named Entity Recognition
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
NLU	Natural Language Understanding
NomBank	Noun Annotation Bank
NPI	Negative Polarity Item
Okapi BM25	Okapi Best Matching 25
OWL	Web Ontology Language
PropBank	Propositional Bank
RDF	Resource Description Framework
RST	Rhetorical Structure Theory
RTE	Recognizing Textual Entailment
S, V and O	Subject, Verb and Object
SDRT	Segmented Discourse Representation Theory
SIGLEX	Special Interest Group on the Lexicon of the Association for Computational Linguistics
SIGNLL	Special Interest Group on Natural Language Learning of the Association for Computational Linguistics
SNLI	Stanford Natural Language Inference
STS	Semantic Textual Similarity

SUMO	Suggested Upper Merged Ontology
SVM	Support Vector Machines
T and H	Text and Hypothesis
TF-IDF	Term Frequency – Inverse Document Frequency
TnT	Trigrams'n'Tags
TREC	Text Retrieval Conference
VENSES	Venice Semantic Evaluation System
WMD	Word Mover's Distance
XML	Extensible Markup Language
YAGO	Yet Another Great Ontology

List of Figures

Page

List of Tables

Page

1 Introduction

1.1 Statement of Problem and Motivation

Please read carefully the following text passage from Daniel Defoe's adventure novel *Robinson Crusoe*[1] (Defoe 2001: 58-59):

> "When I came down from my apartment in the tree I looked about me again, and the first thing I found was the boat, which lay as the wind and the sea had tossed her up upon the land, about two miles on my right hand. I walked as far as I could upon the shore to have got to her; but found a neck or inlet of water between me and the boat, which was about half a mile broad […].
>
> I resolved, if possible, to get to the ship; so I pulled off my clothes, for the weather was hot to extremity, and took the water. But when I came to the ship, my difficulty was still greater to know how to get on board; for as she lay aground, and high out of the water, there was nothing within my reach to lay hold of. I swam round her twice, and the second time I spied a small piece of rope, which I wondered I did not see at first, hang down by the fore-chains so low as that with great difficulty I got hold of it, and by the help of that rope got up into the forecastle of the ship. Here I found that the ship was bulged, and had a great deal of water in her hold, but that she lay so on the side of a bank of hard sand, or rather earth, that her stern lay lifted up upon the bank, and her head low, almost to the water. By this means all her quarter was free, and all that was in that part was dry; for you may be sure my first work was to search and to see what was spoiled and what was free. And first I found that all the ship's provisions were dry and untouched by the water; and being very well disposed to eat, I went to the bread-room and filled my pockets with biscuit, and eat it as I went about other things, for I had no time to lose."

Could you recognize the contradiction in the passage?

The described scene is a topic of many discussions concerning the work of Daniel Defoe (e.g. Baines 2007) and, in general, dealing with logical mistakes occurring in texts of Classic Literature. Robinson Crusoe (referred to by the pronoun *I* during the whole text passage) was intended to get to the wrecked ship. As there was a neck of water between, he had to swim. Before immediately swimming, and as the weather was hot, Robinson took off his clothes. After reaching the ship and being on it for some time, Crusoe found out that some of the provisions had remained dry. He, therefore, went to the bread-room and filled his pockets with biscuits (*I filled my pockets with biscuits*). But how could he do this? Filling the pockets with anything presupposes some clothes with pockets at the time of filling. But as we have been told at the beginning of the text passage, Robinson Crusoe had taken off his clothes before swimming. From reading this at the beginning of the text passage, the reader

[1] The full title of the novel is *The Life and Strange Surprising Adventures of Robinson Crusoe, Of York, Mariner: Who lived Eight and Twenty Years, all alone in an un-inhabited Island on the Coast of America, near the Mouth of the Great River of Oroonoque; Having been cast on Shore by Shipwreck, wherein all the Men perished but himself. With an Account how he was at last as strangely deliver'd by Pyrates*. The novel was first published in 1719.

inferred that Robinson Crusoe did not have pockets during his stay aboard the ship. There was also no information or further clues that Crusoe had found and worn any clothes and put them on. We obviously deal with a contradiction here – two statements express propositions that cannot be true at the same time with the same respect.

It is beyond debate that the recognition of contradictions presents a challenging task for the reader (Markman 1979; Garner 1980, 1981) and especially, for the poor readers (Garner 1980, 1981; Winograd/Johnston 1982). How well an individual performs in detecting contradiction depends on the state of his language and world knowledge, analytical ability, memory as well as his individual characteristics such as, e.g., age (Kotsonis/Patterson 1980; Chan et al. 1987; Vosniadou et al. 1988; Otero/Campanario 1990). The type of contradictions (Markman 1979; Markman/Gorin 1981; Harris et al. 1981; Flavell et al. 1981; Paris/Myers 1981; Garner 1981; Baker 1985) and the preceding notification about the presence of contradictions in the text (Winograd/Johnston 1982; Glenberg et al. 1982; August et al. 1984; Baker/Zimlin 1989) can be crucial for the success of this task as well.

Only a few attempts have been made to reveal and describe the processes involved in the recognition of contradictions by a human. The most prominent theories have been developed by the psychologists in the framework of reading comprehension and described in Otero and Kintsch (1992), Singer (1996), Johnson-Laird et al. (2004) and van den Broek et al. (2005). The proposed theories differ with respect to the model of reading comprehension which they are based upon.

The focus of the present study are contradictions occurring in and between online news texts. There are a number of definitions for contradiction, which, according to Grim (2004), can be grouped into four classes: (1) those which define contradiction in terms of truth and falsity (Prior 1967: 458; Bonevac 1987: 25; Wolfram 1989: 163; Sainsbury 1991: 369) such as in (D1) as follows, (2) in terms of content or form (Reichenbach 1947: 36; Mendelson 1964: 18; Haack 1978: 244; Kalish et al. 1980: 18; Forbes 1994: 102) such as in (D2), (3) in terms of assertion and denial (Strawson 1952, 2011: 16-19; Quine 1959: 9; Brody 1967: 61; Kahane 1995: 308) such as, e.g., in (D3), and (4) as a state of affairs (Routley/Routley 1985: 204) such as, e.g., in (D4). Grim (2004) refers to these four groups as semantic, syntactic, pragmatic, and ontological, respectively.

D1 Two propositions are contradictories if and only if it is logically impossible for both to be true and logically impossible for both to be false. (Sainsbury 1991: 369)

D2 Wff* of the form ‘A & ¬A’; statement of the form ‘A and not A’ (Haack 1978: 244)

D3 A contradiction both makes a claim and denies that very claim. (Kahane 1995: 308)

D4 A contradictory situation is one where both B and ¬B (it is not the case that B) hold for some B. (Routley/Routley 1985: 204)

Though these definitions can be used by humans for recognizing contradictions, they are practically, with the exception of the third group of definitions and only by considering some limitations, only with difficulties applicable for the purpose of the study, which is the development of a system for automatic detection of contradictions in news texts. For instance, no machines are capable of determining the truth value of a sentence at present.

It is obvious that most of the above definitions, to some degree, build on one of the three versions of Aristotle's Law of Non-Contradiction (Section 3.1.1). Thus, the third group of definitions, for example, seems to reflect the ontological version of the law (not to be confused with Grim's ontological definition), which states that "it is impossible that the same thing can at the same time both belong and not belong to the same object and in the same respect, and all other specifications that might be made, let them be added to meet local objections" (*Metaphysics* IV 3 1005b19–23). In our opinion, this formulation is more applicable to development of a system for automatic detection of contradictions and will, therefore, be mentioned prior to the purpose of the study.

It is to note that, besides contradiction, also contrariety will be considered in the present study. Though, both terms will be referred to here as contradiction (compare to German: kontradiktorischer Widerspruch vs. konträrer Widerspruch), they have to be clearly distinguished as not synonymous. The difference between contradiction and contrariety will be presented in Section 3.1.2.

According to the survey on the news consumption across twelve countries conducted in 2015[2] by the Reuters Institute for the Study of Journalism, Oxford University over four channels of news access – television, online (including social media), radio, and printed newspapers, the first two appeared to be the most popular ways of accessing news on a weekly basis, with television being the number-one source in, i.a., Germany (82%), France (80%), and UK (75%), among others, and online access in, i.a., Urban Brazil (91%), Finland (90%), Spain (86%), and Denmark (85%). However, taking into consideration that this survey has been conducted online and thus, may underrepresent users who do not use online services, it can be concluded that TV news is still ahead in the countries that participated in the survey; however, with the clear exception of the United States and possibly Denmark, Finland, and Australia. Moreover, from comparing the news consumption among people of different ages, it can be observed that young people prefer online news and often completely abandon television news. This trend is especially observed for United States, France, and Denmark.

[2] The online report on the survey can be found by following this http://www.digitalnewsreport.org/survey/2015/sources-of-news-2015/

To study in particular online news consumption, the Reuters Institute conducted a survey across 36 countries (i.a. USA, Mexico, Australia, EU countries) in five continents.[3] According to the survey, around a half of the survey participants (54%) across all countries, with a predominance of Southern Europe and Latin America, prefer social media as a source of news in contrast with other sources. However, in Spain, Germany, and France, a reverse or slowing trend for this can be observed. Further, the report shows that 23% use messaging apps (e.g., WhatsApp, Viber, We Chat, FB Messenger, Line, Kakao Talk) for weekly accessing the news. Additionally, it was found out that the access of news via smartphones had increased in comparison to computers and tablets, which amounted to 56%, a score which had doubled since 2013.

With the Internet era, not only the readers' preferences for news source (especially of young readers) have changed. The journalistic practice of news production, i.e. information collection and reporting, has been influenced by the possibilities provided by the Internet as well. A number of studies have been conducted which reveal the changes the Internet had brought to the process of news production, including Reddick and King (2001), Miller (1998), Singer (2003), and Fenton (2012), among others. Fenton (2012) summarizes the research findings, i.a., under the umbrella of criteria such as (data transfer) speed and (web) space.

The great amount of space provided in the web means the production of more news for the journalistic practice. Fenton (2012: 559) frames this as "space equals more news". Space provides a possibility of archiving and updating the news, achieving "more depth of information coverage" (ibid.). Space allows a storage of news in different multimedia formats, and not only as text. Space and speed enable a geographical reach so that journalists do not need to leave their newsroom to write about events that have happened in the world.

Speed enabled by Internet, in turn, for the practice of news production, means an increasing value of immediacy (Fenton 2012). However, while the immediate release and update of the news texts is doubtlessly an advantage for the news reader, it is unfortunately often only possible at the cost of information quality (Gunter 2003; Fenton 2012; Silvia 2001). Taking an advantage of the Internet speed, news organizations often publish their news on the web "before the usual checks for journalistic integrity have taken place" (Fenton 2012: 561). This in turn results in the observation that news texts often include typographical, factual, and logical errors, violating accuracy as one of the fundamental values of news text production, misinforming the reader, and negatively affecting the credibility of the newspaper (Bell 1991; Maier 2005; Bednarek/Caple 2012).

[3] The Digital News Report 2017 on the survey published online can be found by following this link: http://www.digitalnewsreport.org/survey/2017/resources-2017/

Factual errors, according to Silverman (2007), represent the most frequent kind of errors occurring in news texts. In contrast to typographical and logical errors, which can be recognized within the text itself, incorrect facts can be revealed only by applying world knowledge or by referring to the original or other related information sources. Typographical errors, in turn, are not critical and can nowadays be easily recognized by means of autocorrection. In contrast, logical errors, which are the result of a violation of logical laws, e.g., the Law of Non-Contradiction (LNC) and the Law of Excluded Middle (LEM), are the most challenging kind of errors for recognition. In practice, both factual and logical errors, in most cases, remain unnoticed by the reader of the news and are taken for granted as reliable or trustworthy (Svintsov 1979; Bell 1991). The omission of the errors can be a consequence of missing world knowledge required or of a lack of readers' attention while reading. In any case, the reader of the news is misinformed and is not aware of this.

However, if detected by a reader, the typographical, but especially factual and logical errors that have occurred present a negative impact on the newspaper's credibility and trustworthiness since they are perceived as lies or disinformation (Svintsov 1979; Bell 1991; Silverman 2007; Bednarek/Caple 2012). Therefore, in the process of news production, the task of news editing is essential and cannot be ignored. Editing has become even more urgent today because the modern reader has even more possibilities of verifying the information provided, in comparison to the past, as a large amount of related information appears online simultaneously (Silverman 2007).

One should also consider that incorrect facts (factual errors) and logically wrong conclusions (logical errors) in news texts are often used for intentionally serving the purpose of manipulation or propaganda. Violating the news value of objectivity (Section 4.1.3), the facts are adjusted to influence the reader's opinion, forcing it into a particular direction to the advantage of the country's, institution's, or individuals' interests. In this context, in particular, the current phenomenon of fake news reportedly occurring in the social media should be mentioned.

Today in many fields of human life, computers successfully play a supporting role, taking over natural language tasks such as, e.g., searching among a huge amount of data and delivering the needed information in the shortest amount of time, as well as typographical error correction, opinion mining, etc. The main aim of the present study is to propose an approach for automatic detection of contradictions (henceforth referred to as CD) in news texts.

This approach can be of practical relevance first, for the task of news editing when proofing the text for consistency (=agreement with facts previously stated, no contradictions contained). Second, it can be applied to identify on which facts and aspects the different

sources of information disagree and in such a way, to serve the purpose of information verification. Third, an automatic CD task can be used to obtain a summarized view of contradictory opinions and facts on particular events from a large number of news texts in order that a reader can independently form his opinion based on a full picture. Finally, the approach can be integrated into other natural language systems and applications such e.g., question-answering systems and text summarization which among others use news texts as their data source.

From the theoretical perspective, the significance of the study consists first in summarizing and elaborating the existing theoretical knowledge on natural language contradiction. Second, the study provides new empirically gained insights into the realization mechanisms of natural language contradictions occurring in and between news texts, in this way contributing to a better understanding of the nature of contradictions and filling the knowledge gaps.

1.2 Subject of the Study

Natural language contradictions are of complex nature. As will be shown in Chapter 5, the realization of contradictions is not limited to the examples such as *Socrates is a man* and *Socrates is not a man* (under the condition that Socrates refers to the same object in the real world), which is discussed by Aristotle (Section 3.1.1). Empirical evidence (see Chapter 5 for more details) shows that only a few contradictions occurring in the real life are of that explicit (prototypical) kind (see, e.g., Svintsov 1979; de Marneffe et al. 2008). Rather, contradictions make use of a variety of natural language devices such as, e.g., paraphrasing, synonyms and antonyms, passive and active voice, diversity of negation expression, and figurative linguistic means such as idioms, irony, and metaphors. Additionally, the most sophisticated kind of contradictions, the so-called implicit contradictions, can be found only when applying world knowledge and after conducting a sequence of logical operations such as e.g. in (1.1).

(1.1) The first prize was given to the experienced grandmaster L. Stein who, in total, collected ten points (7 wins and 3 draws). (Svintsov 1979: 195)

Those familiar with the chess rules know that a chess player gets one point for winning and zero points for losing the game. In case of a draw, each player gets a half point. Built on this idea and by conducting some simple mathematical operations, we can infer that in the case of 7 wins and 3 draws (the second part of the sentence), a player can only collect 8.5 points and not 10 points. Hence, we observe that there is a contradiction between the first and the second parts of the sentence.

Implicit contradictions will only partially be the subject of the present study, aiming primarily at identifying the realization mechanism and cues (Chapter 5) as well as finding the parts of contradictions by applying the state of the art algorithms for natural language processing without conducting deep meaning processing. Further in focus are the explicit and implicit contradictions that can be detected by means of explicit linguistic, structural, lexical cues, and by conducting some additional processing operations (e.g., counting the sum in order to detect contradictions arising from numerical divergencies).

One should note that an additional complexity in finding contradictions can arise in case parts of the contradictions occur on different levels of realization. Thus, a contradiction can be observed on the word- and phrase-level, such as in a *married bachelor* (for variations of contradictions on lexical level, see Ganeev 2004), on the sentence level – between parts of a sentence or between two or more sentences, or on the text level – between the portions of a text or between the whole texts such as a contradiction between the Bible and the Quran, for example. Only contradictions arising at the level of single sentences occurring in one or more texts, as well as parts of a sentence, will be considered for the purpose of this study. Though the focus of interest will be on single sentences, it will make use of text particularities such as coreference resolution without establishing the referents in the real world.

Finally, another aspect to be considered is that parts of the contradictions are not necessarily to appear at the same time. They can be separated by many years and centuries with or without time expression making their recognition by human and detection by machine challenging. According to Aristotle's ontological version of the LNC (Section 3.1.1), however, the same time reference is required in order for two statements to be judged as a contradiction. Taking this into account, we set the borders for the study by limiting the analyzed textual data thematically (only nine world events) and temporally (three days after the reported event had happened) (Section 5.1). No sophisticated time processing will thus be conducted.

1.3 Research Questions and Objectives

As previously mentioned, the main aim of the present study is to propose a system for automatic detection of naturally occurring contradictions in and between news texts published in English. As regards to the aim of the study, we formulate the following three blocks of related research questions:

RQ1 What conditions must two sentences necessarily satisfy in order to be judged a contradiction? Are there any natural language exceptions?

RQ2 What are the cues of contradictions occurring in news texts written in English? Do all contradictions occur explicitly in news texts?

RQ3 What phenomena of natural languages should a CD system be able to cope with? Considering this, how can the architecture of a system for the automatic detection of contradictions occurring in and between news texts look like? What is the most efficient way of the computational realization of the system's components? What are the current limitations? How can a CD profit from the properties of a text?

The research objectives serving as milestones toward the main aim of the study, are as follows:

O1a Review the state of the art of the CD systems, identify their weaknesses and strengths, and determine the aspects or components that are to be improved;

O1b Review the existing datasets of contradictions and decide whether they can be applied as the basis for the development and evaluation of the CD system. If required, collect and prepare own data;

O2a Based on the existing theory, formulate a set of conditions and rules that underlie the realization of natural language contradictions;

O2b Describe natural language phenomena which can be problematic issues for the CD task;

O3a Outline the characteristics and particularities of the text and, in particular, of the online news texts that have to be considered by a CD system and can potentially contribute to the efficiency of the CD task;

O3b Identify the linguistic cues of naturally occurring news contradictions and offer a typology of contradictions based on these cues;

O4 Develop an architecture of a prototype CD system and implement the system. Decide on which methods and approaches can be used for implementing the system's components and evaluate them on real cases.

1.4 Structure of the Thesis

The overall structure of the study consists of ten chapters, including *Introduction*, *Conclusions*, *References*, and *Appendix*.

After introducing the reader with the subject, the main aim, and goals of the study (Chapter 1, *Introduction*), Chapter 2 (*State of the Art*) begins with the presentation of the main stages of the development of the CD task. It then goes on with an overview of the existing CD systems, summarizes their weaknesses and strengths, and defines the research gaps to be addressed in the study (Section 2.1). Finally, the chapter provides a description of the available datasets of contradictions, which are an essential condition for the development and evaluation of CD systems (Section 2.2).

The next two chapters (together with Chapter 6) lay out the theoretical dimensions of the research, addressing the concepts of contradiction in logic and language (Chapter 3) and the characteristics of news texts with a focus on online news texts (Chapter 4).

In more detail, Chapter 3 (*Contradiction in Logic and Language*), which consists of five sections, is concerned with the traditional approaches to contradiction in logic and language. Section 3.1 first presents the traditional view on contradiction as developed by

Aristotle and then provides a terminological distinction of contradiction to the related concepts such as contrariety, tautology, and paradox. The focus of Section 3.2 is the realization, expression, and interpretation of negation in natural languages with interest in English. The subject of Section 3.3 is the scientific debate on the status of contradiction in the light of phenomena such as presupposition, modality, vagueness, and ambiguity. Further, Section 3.4 provides an overview of existing classifications of textual contradictions, including typologies from educational psychology and computational linguistics. Finally, Section 3.5 concludes the chapter with a summary of the causes and functions of natural language contradictions, claiming that contradictions are not always “bad”.

Chapter 4 (*The Characteristics of News Texts*) introduces the concept of news texts, including the differences between printed and online newspapers, hard and soft news, and values in news production (Section 4.1), description of a news article structure and its main elements (Section 4.2) as well as a discussion of news language particularities (Section 4.3).

Chapters 5, 6 and 7 focus on the conceptual and physical design as well as implementation of the CD system and constitute with Chapters 5 and 7 the empirical part of the present work.

Chapter 5 (*Typology Construction: Types of Contradictions in News Texts*) describes the computationally oriented methodology and reports the results of a corpus-based typology construction of the contradictions occurring in single or multiple news texts.

Chapter 6 (*Conceptual Design of a CD System and Supporting Tools*) in turn addresses a possible conceptual design of a CD system and provides a theoretical background on computational approaches to meaning processing at lexical, morphological, syntax, semantic, pragmatic and discourse levels essential for the support of a CD system (Section 6.1– Section 6.2.3). Approaches to meaning representation are the topic of Section 6.2.4. The chapter then concludes with a presentation of existing computational sources of lexical and world knowledge (Section 6.2.5).

Chapter 7 (*Physical Design of a CD System and Implementation*) then proposes an approach for the CD task, integrating the gained knowledge, and describes the main steps and experiments conducted with an implementation of the system’s components.

Finally, Chapter 8 (*Conclusions*) summarizes the findings and outlines the limitations of the system developed. In respect to these limitations, the areas and tasks for further research are defined.

1.5 How to Read This Thesis

I would like to conclude the introductory chapter with some useful remarks on how to read this thesis by addressing the use of the examples, terminology, and data.

All examples in the thesis are provided with an ID that follows a particular system. Each ID consists of two digits, separated by a point. The first digit indicates the number of the chapter where the example occurs; the second digit indicates the order of the example in this chapter. Examples of contradictions taken from the compiled corpus are additionally provided with an id that indicates where the example can be found in the corpus. The digital version of the corpus is provided on the USB flash drive submitted along with the present work. The digital version of all supplementary materials attached in the present study can be found on the USB flash drive as well.

2 State of the Art

The present chapter serves the purpose of introducing the reader to the state of the art of an automatic task of textual CD. First, it provides an overview and description of existing CD systems and methods (Section 2.1). In order for a reader to form a well-ordered picture about the state of the art, Section 2.1 begins by sketching the main stages of development of interest in automatic detecting textual contradictions before discussing the methods and systems. Due to the relevance, only a selected number of methods and systems will be presented in detail here. The criteria for the selection of the systems and methods were an underlying methodology, performance evaluation scores as well as experts' opinions. The section then concludes with an outline of weak and strong aspects of the systems indicating the research gaps and sets the objectives for the study. Further, in Section 2.2, a description is given of the datasets of contradictions – the so-called corpora – available which are an essential basis for the development and evaluation of the systems. Additionally, the need of collecting own data, despite the existing ones, is explained in this section.

2.1 Methods and Systems

The interest in an automatic CD within the framework of natural language processing (henceforth NLP) and specifically, as a task of natural language understanding (NLU) has its origin in the mid-1990s and is associated with the FraCas project (Cooper et al. 1996). Since then, a number of systems have been proposed, ranging from the simple and robust shallow approaches relying on lexical overlaps and word frequencies to the precise but challenging, deep approaches conducting an advanced semantic interpretation. The best state-of-the-art systems currently achieve approx. 60% accuracy in identifying contradictions that mainly arise from negation and antonyms.

The initial attempts of automatic CD were theoretical and relied on the methodological apparatus of the first-order logic (FOL) (Cooper et al. 1996; Condoravdi et al. 2003). Crouch et al. (2003), in particular, emphasized the potential of sophisticated FOL approaches such as described in Hirst (1991) and Hobbs (1985). However, no practical implementations of logic- or quasi-logic-driven systems have been proposed until the middle of the 2000s. To the first logical and quasi-logical systems count the system described in Tatu and Moldovan (2007), the BLUE system developed by Clark and Harrison (2009), and a hybrid NatLog-system by MacCartney and Manning (2009).

The first CD system implemented that went beyond the FOL was proposed in Harabagiu et al. (2006). The developers relied only on the capability of the machine-learning algorithms

for textual entailment[4] recognition (Section 6.2.2.2) and considered explicit contradiction cues such as negation and antonyms.

A number of systems for CD in English have been developed during the Recognizing Textual Entailment (RTE) challenges in the years 2007-2009 (RTE-3 Extended Task, RTE-4, and RTE-5 challenges).[5] The main requirement for the systems was a classification of the sentence pairs, provided in the three categories of entailment, contradiction, and unknown, the so-called three-way task (Giampiccolo et al. 2007; Voorhees 2008). The RTE systems are presented in Table 1 (RTE-3 Extended Task), Table 2 (RTE-4), and Table 3 (RTE-5). One should note that the systems submitted in the latter RTE challenges by the same authors are, in most cases, improvements on the earlier RTE submissions.

In addition to the RTE systems, a number of standalone systems for different languages have been developed to the present time as well. These include, among others, systems described in Harabagiu et al. (2006), de Marneffe et al. (2008), Ritter et al. (2008), Kim/Zhai (2009), Ennals et al. (2010), Tsytsarau et al. (2010, 2011), Tsytsarau/Palpanas (2011), Pham et al. (2013), Dînşoreanu/Potolea (2013), Lendvai/Reichel (2016) for English; Wartena et al. (2006) for Dutch; Kawahara et al. (2010), Hashimoto et al. (2012), Andrade et al. (2013), Kloetzer et al. (2013), and Takabatake et al. (2015) for Japanese, and Shih et al. (2012) for Chinese. The standalone CD systems for English are summarized in Table 4.

Both RTE and standalone CD systems have been developed for different application purposes, including, e.g., the improvement of textual entailment recognition tasks (the RTE systems), the improvement of text summarization and question-answering systems (e.g., Harabagiu et al. 2006) as well as the detection and summarization of conflicting opinions in social media and other Web 2.0 platforms (e.g., Kim/Zhai 2009; Ennals et al. 2010; Tsytsarau et al. 2010, 2011; Tsytsarau/Palpanas 2011; Dînşoreanu/Potolea 2013; Lendvai/Reichel 2016). The systems follow different often-combined rationales and methodologies, apply a variety of NLP tools, and with the exception of the RTE systems, are evaluated on different datasets, which makes their comparison and generalization challenging. The execution of the same steps for different purposes makes the systems

[4] The term textual entailment is related to the logical entailment but is used in computational linguistics in a looser and more relaxed sense. The organizers of the RTE challenges provide the following definition of textual entailment: "We say that T entails H if, typically, a human reading T would infer that H is most probably true" (Dagan/Glickman 2004: 4). The parts of logical entailment relation premise and conclusion in the framework of RTE refer to text (*T*) and hypothesis (*H*), respectively.

[5] The RTE-1, RTE-2, RTE-3 (Main, but not the Extended Task), RTE-6, and RTE-7 challenges focused on the recognition of entailments only.

generalization difficult as well. Nevertheless, an attempt of systems' comparison is presented in Table 1 – Table 4.

The comparison of the systems reveals that the CD by means of supervised classification is a preferred method, despite the need for a large amount of data for classifier training and model testing. Based on a set of pre-defined features and manually classified (annotated) examples in contradictions and non-contradictions, a classification algorithm searches for patterns in the pre-classified data (training data) and builds a model which, after a test stage, can then be applied to predict any contradictions in a new corpus. For the classification task, a variety of algorithms have been applied, including, among others, maximum entropy in, e.g., de Marneffe et al. (2008), SVM (Vapnik 1995) in, e.g., Malakasiotis and Androutsopoulos (2007), decision trees in, e.g., Hickl et al. (2007), nearest (shrunken) centroids (Tibshirani et al. 2003), and random forest (Breiman 2001) in Lendvai/Reichel (2016). The maximum-entropy algorithm has proved to be most efficient so far. For the application of the classifiers, the WEKA machine-learning tool[6] described, e.g., in Smith/Frank (2016) was preferred.

Concerning pre-defined features, some classification-based systems relied on the degree (or score) of similarity between text and hypothesis sentences (for definition of text and hypothesis see Footnote 4) in tokens, lemma, parts-of-speech, and sentence length (e.g. Malakasiotis/Androutsopoulos 2007; Lendvai/Reichel 2016) computed by multiple similarity measures, without considering any other information. For this task, a number of similarity measures have been applied, including among others, the Levenshtein distance, the Jaro-Winkler distance, the Manhattan distance, the Euclidean distance, the cosine similarity, the n-gram distance, the matching coefficient, the Dice coefficient, as well as the Jaccard coefficient. In general, the results show that although classification based on similarity scores works well for recognizing entailments and neutral cases, CD represents a more complex task (Lendvai/Reichel 2016).

Another group of classification-based systems in turn relied on features which are characteristic for contradiction, including negations, antonyms, numerical mismatches as well as mismatches in grammatical functions and thematic roles (Harabagiu et al. 2006; de Marneffe et al. 2008). In contrast to the simple computation of similarity, the detection of a contradictory relation requires additional steps such as a unified comparable representation of

[6] https://www.cs.waikato.ac.nz/ml/weka/index.html

Feature		Study							
		Malakasiotis/ Androutsopoulos	Clark et al.	Tatu/Moldovan	Hickl et al.	Bobrow et al.	MacCartney/Manning	Iftene/ Balahur-Dobrescu	Wang/Neumann
Accuracy (%)		49.4	45.1	71.3	73.1	43.6	59.1	56.9	45.5
Preprocessing				X	X				
Parsing				X					
SRL				X		X			
Anaphora resolution				X	X				
Lexical resources			X	X	X	X		X	
Paraphrasing			X	X	X			X	
World knowledge			X						
Meaning Representation	Bag-of-words								
	Logical form		X	X			X		
	Dependency graph / tree		X					X	X
	Other								X
Alignment								X	X
Machine learning		X			X				X
String similarity		X			X				
Topic identification									
Contradiction clues	Negation	X	X	X				X	
	Opposition		X						
	Other								
Sentiment analysis									
Datasets	RTE (original)	X	X	X	X	X	X	X	X
	RTE (modified)								
	Other						X		

Table 1: CD Systems submitted for the RTE-3 challenge – Extended Task (2007).

Feature		Study													
		Galanis/Malakasiotis	Clark/Harrison	Glinos	Wang/Neumann	Agichtein et al.	Montalvo-Huhn/Taylor	Varma et al.	Krestel et al.	Siblini/Kosseim	Li et al.	Castillo/Alonso i Alemany	Padó et al.	Iftene	Mohammad et al.
Accuracy (%)		67.6	54.7	41.6	61.4	54.7	46.6	30.9	43.2	61.6	58.8	54.6	55.3	68.5	55.6
Preprocessing					X	X	X	X		X	X	X	X	X	X
Parsing			X	X	X			X	X	X			X	X	X
SRL				X						X	X				
Anaphora resolution				X		X				X			X		X
Lexical resources		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Paraphrasing			X		X										
World knowledge													X		X
Meaning representation	Bag-of-words		X												
	Logical form		X												
	Dependency graph / tree			X	X			X	X	X			X	X	X
	Other														
Alignment				X				X	X	X			X		X
Machine learning		X			X	X				X	X	X	X		X
String similarity		X				X	X					X			
Topic identification													X		
Contradiction clues	Negation	X									X		X	X	X
	Opposition						X						X	X	X
	Other					X					X		X	X	X
Sentiment analysis															
Datasets	RTE (original)	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	RTE (modified)														
	Other														

Table 2: CD systems submitted for the RTE-4 challenge (2008).

Feature		Study									
		Malakasiotis	Clark/Harrison	Han Ren et al.	Wang et al.	Ferrández et al.	Breck	Castillo	Varma et al.	Krestel et al.	Iftene/Moruz
Accuracy (%)		57.5	54.7	52.2	63.7	60	57	52.2	46.9	48.7	68.3
Preprocessing				X		X	X	X	X		X
Parsing		X	X	X	X		X		X	X	X
Semantic Role Labelling					X				X		
Anaphora Resolution				X	X						
Lexical resources		X	X	X	X	X	X	X	X	X	X
Paraphrasing			X				X				X
World knowledge				X		X					
Meaning Representation	Bag-of-words		X								
	Logical form		X								
	Dependency graph / tree				X					X	X
	Other										
Alignment					X					X	
Machine learning		X		X	X	X		X			
String similarity		X		X		X	X	X			
Topic identification											
Contradiction clues	Negation	X	X			X	X		X		X
	Opposition						X		X		X
	Other								X		X
Sentiment analysis											
Datasets	RTE (original)	X	X	X	X	X	X	X	X	X	X
	RTE (modified)										
	Other										

Table 3: CD systems submitted for the RTE-5 challenge (2009).

Feature		Study									
		Harabagiu et al. (2006)	de Marneffe et al. (2008)	Ritter et al. (2008)	Kim/Zhai (2009)	Ennals et al. (2010)	Tsytsarau/Palpanas (2011)	Tsytsarau et al. (2010, 2011)	Pham et al. (2013)	Lendvai/Reichel (2016)	Wartena et al. (2006) for Dutch
Accuracy (%) Precision (%) Recall (%)		64/-/-	-/22.95/19.44	-/62/19	n.a.	n.a.	n.a.	n.a.	-/14/19.44	iPosts -/40/34 Threads -/42/35	n.a.
Preprocessing		X	X	X					X	X	
Parsing		X	X						X		
SRL		X	X						X		
Anaphora resolution		X	X	X					X		
Lexical resources		X	X	X	X				X		
Paraphrasing											
World knowledge			X	X							
Meaning Representation	Bag-of-words				X		X	X	X		
	Logical form										
	Dependency graph / tree	X	X								
	Other			X					X		X
Alignment		X	X	X	X				X		
Machine learning		X	X							X	
String similarity					X					X	
Topic identification			X					X			
Contradiction clues	Negation	X	X		X						
	Opposition	X	X		X				X		
	Other	X	X	X					X		
Sentiment analysis					X		X	X			
Datasets	RTE (original)		X						X		
	RTE (modified)	X	X								
	Other		X	X		X	X	X		X	X

Table 4: Standalone CD systems.

text and hypothesis meaning and their alignment. The preferable means for meaning representation were dependency trees converted to typed dependency graphs, e.g., in de Marneffe et al. (2008), functional dependency triples alone (Wang/Neumann 2008) or combined with frame representation based on semantic role frames (Pham et al. 2013), the functional dependency tuple (Ritter et al. 2008) as well as the bag-of-words (Tsytsarau et al. 2010, 2011; Tsytsarau/Palpanas 2011), only to name a few. For the representation of sentences as a functional dependency of a verb predicate and two arguments, the REVERB tool (Fader et al. 2011) applied in Pham et al. (2013) and the TextRunner Open Information Extraction system (Banko et al. 2007; Banko/Etzioni 2008) has been applied in Ritter et al. (2008). For alignment, despite a greedy algorithm, the maximum entropy-based classifier was preferred (Hickl et al. 2006) in, e.g., Harabagiu et al. (2006).

In addition to classification- and rule-based systems, the third group of systems adopt a slightly loose logical form in their meaning representation and incorporate logical inference rules (Tatu/Moldovan 2007; Clark and Harrison 2009; MacCartney/Manning 2007) as well as detect contradictions based on opposite sentiments and statistical computing (Tsytsarau et al. 2010, 2011; Tsytsarau/Palpanas 2011; Dînşoreanu/Potolea 2013) or patterns over ontology terms (Wartena et al. 2006).

Common for all systems is the use of lexical resources (Section 6.2.5.1) such as WordNet (Fellbaum 1998), VerbNet (Kipper et al. 2000), and DIRT (Lin and Pantel 2001) for identifying meaning relations (i.a., oppositions and synonyms) for the purpose of sentence alignment, improving the building of a classification model and detecting contradictions. For knowledge-based contradictions, the Wikipedia resource was most preferred.

A number of studies emphasize the importance of finding related text and hypothesis sentences which describe the same event in order to achieve better performance of the systems on CD task (de Marneffe et al. 2008; Kim/Zhai 2009; Pham et al. 2013; Lendvai/Reichel 2016). The authors proceed on the assumption that two events cannot be contradictory when they are not related. The related sentences were found in the proposed systems by means of, e.g., a Jaccard similarity function in combination with WordNet by, e.g., Kim/Zhai (2009) as well as a latent Dirichlet allocation (LDA) topic modelling algorithm (Blei et al. 2003) at a sentence level (Denecke/Brosowski 2010) applied in Tstytsarau et al. (2011).

The general natural processing tasks integrated into the systems include data normalization (i.a., temporal, abbreviations, etc.), parsing for the purpose of identifying grammatical functions and constructing meaning representations, part-of-speech tagging, anaphora resolution within a sentence or between two neighbor sentences, semantic role labeling for identifying the thematic roles, polarity computing, and others. For parsing, the Charniak parser

(Charniak 2000), chart parser SAPIR (Harrison/Maxwell 1986), Collins parser (Collins 2003), Stanford dependency parser (Klein/Manning 2003; de Marneffe et al. 2006) and MiniPar (Lin 1994) have been applied. The LingPipe tool (e.g., described in Baldwin/Dayanidhi 2014) was a preferred toolkit for named entity recognition (NER) and TnT (Brants 2000) for part-of-speech tagging. Anaphora resolution in turn has been performed, e.g., by means of a tool which combines the Hobbs algorithm (Hobbs 1978) and the resolution of anaphora procedure (Lappin/Leass 1994). Semantic role labeling was conducted by means of, e.g., the SENNA package (Collobert et al. 2011). For normalization of time expressions, e.g. the TARSQI toolkit (Verhagen et al. 2008) has been applied. Only a few systems (Harabagiu et al. 2006; de Marneffe et al. 2008) make use of information on modality and quantification, which is essential for the task of CD.

To the most prominent, most cited, and interesting CD approaches for English belong to those developed and described in Harabagiu et al. (2006) and de Marneffe et al. (2008), as well as its improvement and extension proposed in Padó et al. (2008) and Ritter et al. (2008), and sentiment-based CD presented in Tsytsarau et al. (2010, 2011) and Tsytsarau/Palpanas (2011).

As already mentioned earlier, Harabagiu et al. (2006) were the first to provide empirical results for the task of CD. The authors point out that the task can increase the quality of other NLP tasks such as question-answering and multi-document summarization. In the case of discovering contradictory information from multiple sources, the systems have to decide which information is preferred for the output. For this, the inconsistent information can either be checked by the additional intervention of a user or by contacting additional knowledge resources.

The system proposed in Harabagiu et al. (2006) detects contradictions by following two views. According to the first view, contradictions can be recognized by removing the negations of propositions (argument-predicate structure) and then testing the propositions for textual entailment. Harabagiu et al. (2006) used their own textual entailment system for conducting this task. According to the second view, contradictions can be detected by training a classifier upon positive representatives of the contradictions relying on linguistic information such as negations (*n't*, *not*; verbs *to deny*, *to fail*; prepositions *without*, *except*, etc.), antonyms as well as explicit cues of contrast relations (e.g., *but*, *although*, *however*). For the classification task, the maximum entropy machine learning algorithm was applied.

To train and evaluate the classifier for detecting the contradictions arising from negations and antonyms, a modified RTE-2 dataset (for more information, see Section 2.2.2) has been used. For training and evaluation of the classifier for recognizing contrast relations datasets

of a total of 10,000 sentence pairs (9,000 training datasets and 1,000 evaluation datasets) have been collected from online news articles.

The results of the training and the following testing of the system showed that the system, by following the second view, shows better performance in CD. The proposed approach could achieve a 62% overall accuracy in identifying contradictions arising from negation and antonyms.

A similar but more extended system was proposed in de Marneffe et al. (2008). Analogous to Harabagiu et al. (2006), the system makes use of the predicate-argument, meaning representation, recognition of textual entailment, and supervised machine learning techniques but relies, in contrast, the system of Harabagiu et al. (2006) not only on information of negation and antonyms.

Moreover, the authors compiled the first corpus of naturally occurring contradictions, representing a more realistic data basis for system development (Section 2.2.3). Based on their corpus, de Marneffe et al. (2008) constructed a typology of contradiction cues, including negation, antonymy, numerical mismatches, structural, factivity, and modality information as well as world knowledge (see Section 3.4.3.2 for more information on these types). The authors point out that the contradictions arising from the first three features are relatively easy to model and detect as no deep comprehension is required. Detecting the contradictions marked by the latter aspects, in turn, requires a more precise meaning modeling.

The system proposed in de Marneffe et al. (2008) is based on the Stanford RTE system (MacCartney et al. 2006) and was extended by an additional step of event coreference recognition. The authors claim that sentences about different events cannot be contradictory. However, as the result of missing context, sentences such as (2.1) were assumed to be contradictory without further analyzing whether *woman* refers to the same person.

(2.1) Passions surrounding Germany's final match turned violent when a *woman* stabbed her partner because she didn't want to watch the game.

A *woman* passionately wanted to watch the game.

In general, the CD process by the Stanford system consists of four steps. First, the input text and hypothesis sentences are syntactically and semantically analyzed by means of the Stanford dependency parser (Klein/Manning 2003; de Marneffe et al. 2006) and then converted to typed dependency graphs. In the second step, based on the similarity and syntactic information that was combined by means of the margin infused relaxed algorithm (Crammer/Singer 2001), the graphs are aligned with each other, if possible. Padó et al. (2008) offered an improvement on this step by applying the edit distance-based alignment system MANLI (MacCartney et al. 2008) and the stochastic aligner. In the third step, sentences that are not related and do not describe the same event are filtered out by the system. Two

different approaches have been proposed for this task. The authors claim that on one side, the root of the hypothesis graph aligned with text graph can indicate the co-referent events. It is, however, efficient in the case when the hypothesis sentences are shorter than the text sentences. On the other side, the authors propose modeling the sentence topicality as a technique for co-referent event detection. The two approaches were tested on the RTE-3 development dataset. The results are presented in Table 5, indicating first that the two approaches, in general, have to be improved and second, addressing the need for other techniques for filtering the non-co-referent events. Finally, in the fourth step, the contradictory features are extracted, and logistic regression is applied to classify the hypothesis and text sentences as contradictory or not.

Approach	**Precision**	**Recall**
No filter	55.10	32.93
Root alignment	61.36	32.93
Root alignment + topicality modelling	61.90	31.71

Table 5: Comparison of approaches to graph alignment applied in de Marneffe et al. (2008).

To test the system, the modified RTE-1_test, the RTE-2_test (contradictions arising from negations) dataset, and the original RTE-3_test dataset were used. The authors report a 42.22% precision and a 26.21% recall for detecting contradictions in the RTE-1_test dataset, a 22.95% precision and a 19.44% recall for the RTE-3_test dataset, and a 62.97% precision, a 62.50% recall, and a 62.74% accuracy for the modified RTE-2_test dataset of negation. Further, the comparison of the results for each contradiction type separately shows that the system is efficient in the detection of contradictions arising from negation, antonyms, and numeric mismatches and needs improvement in the detection of lexical and world knowledge contradictions.

Ritter et al. (2008) proposed an extension of the Stanford system, addressing the problem of world knowledge contradictions in their study, such as in (2.2):

(2.2) a. Mozart was born in Salzburg.

b. Mozart was born in Vienna.

Here, a contradiction arises as the result of the incompatibility between *Salzburg* and *Vienna* in respect to the co-referent subject *Mozart* and driven by the relation expression *was_born_in*. This kind of relation, which can be formally represented as *R(x,y)*, the authors call *functional*. The functional relation e.g. in (2.2a) can thus be represented as *was_born_in(Mozart, Salzburg)*. The relation *R* between *x* (subject) and *y* (object) is a functional relation if and only if *x* is not ambiguous and is not related to different entities in the real world, and the function *R* maps *x* to the unique variable *y*.

For the detection of contradictions marked by functional relations, Ritter et al. (2008) proposed a three-staged domain-independent system which they called AuContraire. In the first stage, the system analyzes sentences and presents them as one or more tuples that have the form *R(x,y)*. For this task, the TextRunner component of the Open Information Extraction system (Banko et al. 2007; Banko/Etzioni 2008) was applied. In the second stage, the system identifies pairs of sentences which, with high probability, are functional relations and groups them into a set of *R(x, •)* with the same subject. For this, the authors propose the application of a modified expectation-maximization algorithm (Dempster et al. 1997). Finally, in the third stage, the system filters out cases, such as in (2.3), by reasoning about the synonymy, meronymy, and type of *x* and *y* (person, data, location, etc.) and identifying the non-co-referent arguments. For identifying the meronyms, the developers used diverse lexical resources such as Tipster Gazetteer and WordNet (Fellbaum 1998). The synonyms, in turn, were recognized by computing the edit distance and string similarity (Cohen et al. 2003), as well as by applying a RESOLVER system for synonyms identification (Yates/Etzioni 2007), and by WordNet. To identify the type of x and y, the task of NER was performed in combination with lists of personal and geographical names.

(2.3) Alan Turing was born in London.

Alan Turing was born in England.

To evaluate the system, Ritter et al. (2008) first used TextRunner to collect 1000 relations automatically from 117 million Web pages. They labeled each relation as a functional or a non-functional relation. They achieved a 62% precision and a 12% recall, and a 92% recall and a 51% precision on the balanced data (contradictions and non-contradictions in a proportion of 1:1).

2.2 Corpora of Contradictions

2.2.1 FraCas Inference Data Suite

The FraCas (A Framework for Computational Semantics) inference test suite is considered to be the first corpus for English, which includes contradictions together with examples of entailments. The dataset was developed within the scope of a joint project of the Universität des Saarlandes (Germany), Universität Stuttgart (Germany), and University of Edinburgh (United Kingdom) in the middle of the 1990s (Cooper et al. 1996). The purpose of the project was to provide data for the development, evaluation, and improvement of applications for NLP focusing on inference processing. Cooper et al. (1996) define the central capability of such applications to be the ability of inference processing.

The FraCas corpus consists of 346 units, so-called problems, each including 1-5 statements (premises), one yes/no question, and a yes/no answer, where *yes* indicates an entailment,

no a contradiction, and *don't know* remains for neutral cases. Some *yes* and *no* answers additionally include comments and explanations as in the example of (2.4). The number of premises in the problems in total amounts to 536. The distribution of the premises and answers in the corpus is presented in Table 6 and Table 7, respectively.

(2.4) Premise: Dumbo is a large animal.
Question: Is Dumbo a small animal?
Answer: [No]
Large(N) => ¬Small(N)

Number of Premises	Number of Problems	Number of Problems (%)
1	192	55.5
2	122	35.3
3	29	8.4
4	2	0.6
5	1	0.3

Table 6: The distribution of premises in the FraCas corpus.

Answer	Number of Answers	Number of Answers (%)
Yes	180	52
Don't	94	27
No	31	9
Other/complex	41	12

Table 7: The distribution of answers in the FraCas corpus.

In general, the FraCas problems are divided into nine groups, according to the categories involved in semantic inference construction such as quantifiers, plurals, anaphora, ellipsis, adjectives, comparatives, temporal reference, verbs, and attitudes. The problems in each group, in turn, are further divided into subgroups representing single aspects of each category. The problem unit in (2.4) is an example of the category adjectives, the subcategory opposites. The distribution of the problems in the groups is presented in Table 8.

Group of Problems	Number of Problems	Number of Problems (%)
Quantifiers	80	23
Plurals	33	10
Anaphora	28	8
Ellipsis	55	16
Adjectives	23	7
Comparatives	31	9
Temporal	75	22
Verbs	8	2
Attitudes	13	4

Table 8: The distribution of the problems per group in the FraCas corpus.

In 2009, MacCartney improved the FraCas corpus for the purpose of his study and annotated it with XML. Besides conducting some corrections and adding relevant notes, MacCartney (2009) rephrased the questions into declarative sentences, facilitating them for automatic processing. The original version of the FraCas corpus as ps-file and its improved XML-version are freely available for download at the webpage of Stanford University.[7]

2.2.2 RTE Datasets and Their Modifications

A number of datasets, including contradictions, have been developed within the RTE challenge during the period of 2006 to 2011. The RTE datasets were created with the aim of providing a comparable basis for evaluation of the systems participating in the RTE challenges. All datasets are divided into development and text datasets and include mainly manually constructed pairs of sentences, representing entailments and non-entailments (contradictions and neutral cases). The statistics on RTE datasets, partially adapted from Bentivogli et al. (2009), are presented in Table 9.

All RTE datasets are freely available on the web, directly or upon request.[8] Since RTE-6 (Bentivogli et al. 2010) and RTE-7 (Bentivogli et al. 2011) include no annotations of contradictions, as well as no extensions of the datasets as regards to contradictions, the datasets will not any further be taken into consideration.

The RTE-1 (Dagan et al. 2006), RTE-2 (Bar-Haim et al. 2006), and RTE-3 Main Task (Giampiccolo et al. 2007) challenges were interested only in the task of automatic classification of the data in entailments and non-entailments. For this reason, the correspondent datasets are annotated exclusively with the categories entailments (label yes) and non-entailments

[7] https://nlp.stanford.edu/~wcmac/downloads/
[8] https://tac.nist.gov//

(label no), without further specification of non-entailments in contradictions and neutral cases. In terms of the RTE challenge, this is called a two-way task. The three-way task annotation of the RTE-1 and RTE-2 datasets, representing entailments, contradictions, and neutral cases, was later performed by Harabagiu et al. (2006) and de Marneffe et al. (2008).

Challenge	Dataset	Size (No. of pairs)	Hypothesis length (No. of words)	Text length (No. of words)	Contradictions (%)
RTE-1	Dev	567	10.08	24.78	-
	Test	800	10.8	26.04	-
RTE-2	Dev	800	9.65	27.15	-
	Test	800	8.39	28.37	-
RTE-3 (Extended)	Dev	800	8.46	34.98	10
	Test	800	7.87	30.06	9
RTE-4	Test	1,000	7.7	40.15	15
RTE-5	Dev	600	7.79	99.49	15
	Test	600	7.92	99.41	15

Table 9: The statistics on RTE datasets partially adapted from Bentivogli et al. (2009).

Harabagiu et al. (2006) modified the RTE-2 dataset for the purpose of training and testing their system for the detection of contradictions marked by explicit negations (e.g., *not*), antonymy, and contrast discourse relation cues (e.g., *but*, *although*). To our current knowledge, the modified corpus is not available, neither for free nor for purchase.

In modifying the RTE-2 dataset, Harabagiu et al. (2006) followed three different approaches. First, 800 instances of positive entailments from the RTE-2 dataset were manually negated by human annotators, such as shown in (2.5). As a result, a balanced corpus of 800 contradictions (Dataset 1) has been created. In order to avoid overtraining the model, the annotators were also asked to negate 800 examples of negative entailments (=non-entailments) from the RTE-2 dataset. The produced instances (Dataset 2) were then checked to remove contradictions.

(2.5) a. Former dissident John Bok, *who has been on a hunger strike since Monday*, says he wants to increase pressure on Stanislav Gross to resign as prime minister.

b. A hunger strike was *not* attempted.

Second, the human annotators were asked to paraphrase the negative sentences created in each pair of the Dataset 1, such as in the example of (2.6). As the paraphrasing was not possible for all cases, the corpus of 638 out of 800 instances could be created.

(2.6) a. Former dissident John Bok, *who has been on a hunger strike since Monday*, says he wants to increase pressure on Stanislav Gross to resign as prime minister.

b. A hunger strike was *called off*.

Finally, the third dataset was created by combining 800 examples of non-contradictions with a randomly chosen 400 contradictions from the first and second datasets.

Two years later, de Marneffe et al. (2008) proposed modifications and extensions of the RTE-1, RTE-2 and RTE-3 datasets.[9] First, following the methodology of Harabagiu et al. (2006), they modified the RTE-2 dataset by randomly choosing 102 pairs of sentences (51 entailment and 51 non-entailments) from the RTE-2 test dataset and changing them by adding explicit negation. Afterward, they labeled the sentence pairs with yes for contradiction and no for a non-contradiction. The datasets can be downloaded from the website of the Stanford NLP Group[10].

Second, de Marneffe et al. (2008) extended the annotation of the sentence pairs of the RTE-1, RTE-2, and RTE-3 (Main Task) datasets from two-way task labels (*yes* for entailment relation between the sentences in the pair and *no* for non-entailment) to three-way task labels (*yes* for entailment relation between the sentences in the pair, *no* for contradiction, and *unknown* for non-entailment relation, excluding contradiction). For this, each instance of non-entailments in the RTE-1, RTE-2, and RTE-3 datasets was checked whether it is a contradiction or not. The decision is made by following the guidelines prepared by the Stanford project team.[11] The pairs were labeled manually, either by one or two annotators. Moreover, the contradictions in the RTE-1, RTE-2, and RTE-3 datasets were assigned a type of contradiction based on a contradiction type (e.g., negation, antonymy, world knowledge, etc.). More details on the characteristics of each contradiction type are provided in Section 3.4.3.2 of the present work. The distribution of contradictions in the RTE-1, RTE-2, and RTE 3 tests and development datasets is presented in Table 10. According to the statistics, contradictions constitute in total only 10% of the instances in all three RTE datasets. The distribution of contradictions according to their types on the example of the RTE-3 development dataset is presented in Table 11.

Challenge	Dataset	Original file name	Number of contradictions	Total number of instances
RTE-1	development (1)	RTE1_dev1	48	287
	development (2)	RTE1_dev2	55	280
	test	RTE1_test	149	800
RTE-2	development	RTE2_dev	11	800
RTE-3	development	RTE3_dev	80	800
	test	RTE3_test	72	800

Table 10: Number of contradictions in the RTE-1, RTE-2, and RTE-3 datasets.

[9] De Marneffe et al. (2008) explain the need to again modify the datasets by the fact, that the corpora could not be made available by Harabagiu et al.
[10] https://nlp.stanford.edu/projects/contradiction/
[11] https://nlp.stanford.edu/projects/contradiction/contradiction_guidelines.pdf

Type of contradiction	Distribution (%)
Antonym	15.0
Negation	8.8
Numeric	8.8
Factive/Modal	5.0
Structure	16.3
Lexical	18.8
World Knowledge	27.5

Table 11: Distribution of contradictions occurring in the RTE-3 development dataset according to the contradiction type.

Since 2008 three-way task labeled RTE-4 (Giampiccolo et al. 2008) and RTE-5 (Bentivogli et al. 2009) datasets specifying non-entailments into contradiction and unknown have been created. Sentence pairs in the datasets are labeled with *yes* for positive entailment, *no* for contradiction and *unknown* for neutral cases. The methodology of datasets compilation and annotation is the same as for the RTE-2 and is described in more detail in Dagan et al. (2009). The distribution of contradictions in the RTE-4 and RTE-5 datasets (test and development) is presented in Table 9. The main particularity of the RTE-5 dataset toward the other RTE datasets is the larger size of texts, in such a way providing a more realistic data basis for the development and evaluation of CD and RTE systems.

2.2.3 Stanford Corpus of Real-Life Contradictions

Besides modifying and extending the RTE datasets, de Marneffe et al. (2008) additionally compiled a corpus of natural, or “real-life”, contradictions. The authors argue that manually created contradictions from the RTE 1-3 datasets do not necessarily cover the diversity of contradictions naturally occurring in the language and, therefore, provide an insufficient data basis for the development of efficient and effective systems for CD. Additionally, they claim that real contradictions can be more challenging for automatic recognition than the manually created ones.

To compile a corpus of naturally occurring contradictions, de Marneffe et al. (2008) collected 131 pairs of contradictory sentences from the web. The instances included 19 contradictions from news articles (predominately from Google News), 51 from Wikipedia, 10 from the Lexis Nexis database, and 51 from the LDC project data. The sentence pairs were then manually annotated by two annotators with contradiction types. In case of divergences in annotator’s judgments, these have been clarified by discussion with agreement achieved if possible. Unfortunately, no information on an agreement between annotators on contradiction types has been provided by the researchers. The distribution of contradictions according to their type is presented in Table 12.

Type of contradiction	Distribution (%)
Antonym	9.0
Negation	17.6
Numeric	29.0
Factive / Modal	6.9
Structure	3.1
Lexical	21.4
World Knowledge	13.0

Table 12: Distribution of contradictions occurring in the Stanford Corpus of Real-Life Contradictions according to the contradiction type.

2.2.4 SNLI Corpus

Another corpus developed by the Stanford group, not only for the study of contradiction and textual entailment but also for the development of other applications for NLP is the SNLI 1.0 (Stanford Natural Language Inference) balanced corpus. Currently, the SNLI is considered as the largest state-of-the-art corpus for the task of RTE (also natural inference).

The corpus is divided into development, test, and training datasets and consists of a total 570,152 sentence pairs, including examples of entailment, contradiction, and neutral cases. Their distribution in each dataset is presented in Table 13. The total number of instances in the corpus amounts 37,026.

Dataset/Characteristics	Size (No. of pairs)	No. of contradiction	No. of entailments	No. of neutral cases	No. of unlabelled cases
Development	10,000	3,278	3,329	3,235	158
Test	10,000	3,237	3,368	3,219	176
Training	550,152	183,187	183,416	182,764	785

Table 13: Distribution of contradictions, entailments, neutral, and unlabeled cases in the SNLI corpus.

The sentence pairs for the corpus have been created manually in “a grounded naturalistic context” (Bowman et al. 2015: 1) by about 2,500 participants of the crowdsourcing Internet marketplace Amazon Mechanical Turk. For this purpose, the Stanford team developed the following methodology. Each MTurk worker was presented with a caption of a photo that served as a premise and was given a task to write three kinds of hypotheses for this caption, representing entailment (*definitely a true description of the photo caption*), contradiction (*definitely a false description as of a photo*) and a neutral sentence (*might be a true description of a caption of a photo*) for one premise. Photo captions were provided by the Flickr corpus which consists of 160,000 unattributed captions to 30,000 scenes (Young et al. 2014).

Thus, for example, for a caption of a photo *Two dogs are running through a field*, the entailment could be as shown in (2.7a), the neutral sentence as in (2.7b), and the contradiction as in (2.7c). The examples are taken from Bowman et al. (2015: 3).

(2.7) a. There are animals outdoors.

b. Some puppies are running to catch a stick.

c. The pets are sitting on a couch. (Under assumption that both refer to the same point in the time)

In total, 570,152 sentence pairs have been collected. These are presented as original sentences, as syntactically parsed, and as S-ROOT parsed. The premise sentences are predominantly longer than the hypothesis sentences. That is, the mean length of premise sentence is 14.1 tokens, and the mean length of the hypothesis is 8.3 tokens. Moreover, premise and hypothesis are in, most cases, syntactically different from each other. Further, the data in the corpus is not cleaned and includes few mistakes. The SNLI is released under a Creative Commons Attribution-Share Alike 4.0 International License and can be downloaded freely.[12] It is available in the JSON format and as text files with tab separated values.

2.3 Summary

To sum up, the present methods and systems for CD task show good but still insufficient performance. That is, the mean accuracy score that the current systems could achieve accounts for 60%. The relatively low performance of the systems can be explained by the complexity of natural language contradictions, as well as by the diversity of ways and mechanisms of their realization, making the task of automatic CD challenging. The specific reasons for the low performance of the systems can be the following. First, most of the methods were initially developed and tested on the basis of artificially synthesized pairs of contradictory sentences and are, therefore, probably not able to cover the whole diversity of naturally occurring contradictions. Second, the systems developed focus mainly on detection of explicitly expressed contradictions, relying on linguistic features such as negation and antonyms. Only a few methods address the detection of implicit contradictions, which requires more sophisticated processing than the detection of explicitly expressed contradictions. Third, the pairs of contradictory sentences were analyzed out of the context in which they occur, in this way losing helpful information for CD such as e.g. the aspect of coreference between entities and events. Thus, there still remains a need for an efficient method for an

[12] nlp.stanford.edu/projects/snli/

automatic CD indicating, foremost, gaps in the efficient methods for finding related sentences that may potentially form a contradictory or contrary relation.

Though different approaches have been applied to the collection of contradictions, including manual construction and free collection from the web, the manual construction of contradictions has been preferred so far. In our opinion, however, the manually constructed examples do not have a claim to cover the diversity of the naturally occurring contradictions. Additionally, due to the limitations of the manual data creation, contradiction pairs are presented isolated from their text and context, thereby losing valuable information such as, e.g., the co-references (without knowledge about the referents in the real world) that can contribute to the better performance of the systems. Finally, with the exception of corpora that include some single examples, there is no special corpus for news text contradictions. Therefore, built on this background, there arises the need of collecting our own data – contradictions that occur in news texts – for the purpose of the study. Our methodology for collection of contradictions naturally occurring in news texts, along with text they appear in, will be provided in Chapter 5.

3 Contradiction in Logic and Language

The purpose of Chapter 3 is to provide a theoretical background on logical contradiction as one of the key concepts of the study as well as to outline the conceptual direction for the development of a CD system. The chapter begins with a definition of contradiction in logic, drawing upon the theory elaborated by Aristotle in his works *De Interpretatione* (On Interpretation), *Categoriae* (Categories), and *Metaphysica* (Metaphysics) and incorporating some related notions from the classical and non-classical logic (Section 3.1.1). Three views of Aristotle on contradiction – an ontological, a logical, and a psychological – are the topic of the section, whereas the ontological view on contradiction will be of primary interest. Section 3.1.2 introduces the concept of contrariety, which is often confused with contradiction. Due to the close similarity between the concepts, both contradiction and contrariety will be considered for the purpose of the study. Contradiction and quantification are the topics of Section 3.1.3. Further related concepts, such as tautology and paradox, are the topic of Section 3.1.4. Sections 3.2 and 3.3 in turn discuss the status of contradiction in natural languages. Section 3.2 deals with the diversity and realization of negation in natural languages with a focus on English. In particular, it will be shown that negation in natural languages can be realized in multiple ways rather than only by *not* (Section 3.2.1). The aspects of negation, such as word order and scope of negation, are the topics of Sections 3.2.2 and 3.2.3, respectively. Finally, the phenomenon of double and multiple negation will be addressed in Section 3.2.4. The topic of Section 3.3, in turn, is the status and interpretation of contradiction when considering the aspects of presupposition (Section 3.3.1), modality (Section 3.3.2), vagueness (Section 3.3.3) as well as the cancellability of contradiction under context (Section 3.3.4). These aspects are of interest because they partially or completely fall outside the scope of Aristotle's definition of contradiction but notwithstanding, have to be considered for development of an efficient and effective CD-system. An overview of classifications of contradiction in text linguistics, psychological education, and computational linguistics is provided in Section 3.4. Finally, the chapter concludes by outlining the aspects that cause the formation of contradictions and further discusses the functions of contradictions (Section 3.5).

3.1 Contradiction as Concept of Logic

3.1.1 Contradiction in Traditional (Aristotelian) Logic

The idea of contradiction has been introduced into philosophy by Heraclitus and was further elaborated in the works of Parmenides and Plato. However, the first considerable work on nature of contradiction is attributed to Aristotle. Aristotle treated contradiction as, in his terms, "the most certain among all principles" within the LNC[13], claiming that without the LNC, we could not know anything that we know. For an elaborated overview of the historical development of the concept contradiction, see Ganeev (2004). Different approaches to the treatment of contradiction and its related concepts, such as contrariety and subcontrariety

[13] The Law of Non-Contradiction is also referred to in theory as the Law of Contradiction, the Principle of Non-Contradiction, and the Principle of Contradiction.

in traditional (Aristotelean), classical, and non-classical (paraconsistent) logic, are summarized in Béziau (2006).

The early discussions of Aristotle on the LNC take place in *De Interpretatione* and in *Posterior Analytics* I 11. The treatment of the law in *Metaphysics* IV 3–6 is regarded as foundational. According to logicians (e.g., Horn 2014; Gottlieb 2015), Aristotle provides three views (or versions) on the LNC: an ontological, a doxastic, and a semantic view (Gottlieb 2015). These versions are also referred to as ontological, logical, and psychological (Horn 2014), respectively. The first version refers to the things that exist in the world, the second is about our beliefs, and the third one is about the truth value of assertions.

The ontological version is considered to be the main version of the LNC. Aristotle formulates it as follows: "It is impossible that the same thing can at the same time both belong and not belong to the same object and in the same respect, and all other specifications that might be made, let them be added to meet local objections" (*Metaphysics* IV 3 1005b19–23). One should note that the "same thing" is meant to refer to the same object in the real world and not just to a linguistic expression.

According to the second doxastic version of the LNC, "it is impossible for anyone to believe that the same thing is and is not, as some consider Heraclitus said" (*Metaphysics* IV 3 1005b22–25). As Gottlieb (2015) claims, the doxastic version of the law can be regarded as "implausible" with respect to human psychology as people often have inconsistent beliefs. This is especially the case when they take the consequences of their beliefs into account. That is, we often utter something and simultaneously, we do not believe what we are saying. There is no need to believe that something is, saying that it is.

In general, it is not completely clear how Aristotle understands the doxastic version of the LNC. At the end of *Metaphysics* IV 3, he says that doxastic version is based on the ontological one, but instead of operating with the initially formulated belief that not p, he confuses the reader by speaking about not having the belief that p (Gottlieb 2015).

Finally, the third semantic version of the LNC states that "the opinion that opposite assertions are not simultaneously true is the firmest of all" (*Metaphysics* IV 6 1011b13–14). As Gottlieb (2015) argues, this version is neutral with regard to the internal structure of the assertion, focusing only on the truth value of the assertions. However, she further adds that Aristotle assumes that any assertion involves predicating each other. For this reason, the semantic version of the LNC can better be interpreted as the variant of the ontological version.

Which version of the LNC Aristotle considered the main one is a "matter of controversy" (Gottlieb 2015). As, e.g., Wedin (2004) and Horn (2014) argue that it is the ontological version which is primary to the doxastic and semantic versions. Gottlieb (2015), in turn, does

not tie herself to a particular opinion, claiming that it can be the ontological version that is primary to the doxastic and semantic versions or vice versa, or the doxastic version primary to the ontological. In the present study, these are the ontological version foremost, as well as the semantic version, that are the basis for the development of a CD system.

3.1.2 Contradiction and Contrariety

In the literature, the term of contradiction is often used to refer to contrariety. Though contradiction and contrariety are closely related concepts, and their synonymous use can be to some degree justified, they have clearly to be distinguished from each other.

In order to establish the relationship between the two concepts, another indemonstrable logical law addressed by Aristotle – the LEM – has to be introduced. As in case of the LNC, Aristotle provides several definitions of the LEM. That is, in his *Metaphysics* 1011b23, he defines the LEM as “nor, on the other hand, is it possible that there should be anything in the middle of a contradiction, but it is necessary either to assert or to deny any one thing of one thing”. Another view on the LEM, which is truth value-oriented, can be found in Aristotle’s *De Interpretatione* 18a31: “And with universals taken universally it is always necessary for one to be true and the other false, and with particulars too, as we have said”.

So, how are these two laws related to contradiction and contrariety? According to Aristotle, both laws apply to contradictions, while contrariety arises only in the case the LNC is obeyed. “Nothing can exist between two contradictories, but something may exist between contraries.” (*Metaphysics* 1055b2) That is, “a dog cannot be both black and white, but it may be neither” which is a contrariety. In contrast, in the case of the contradiction illustrated in (3.1), one can be either dead or alive.

Thus if, e.g., regarding the truth-value-oriented versions of the LNC and LEM, a contradiction is present if the affirmation is true when the negation is false and vice versa, the negation is true whenever the affirmation is false, such as stated in the LNC, and the affirmation and the negation cannot at the same time both be false or true, as stated in the LEM.

(3.1) Socrates is dead.

Socrates is alive.

It is to note, that in the case of both contradiction and contrariety, their parts cannot be simultaneously true. If it is the case, then we are dealing with a tautology (Section 3.1.4).

3.1.3 Contradiction and Quantification

The relation between contradiction and quantification is traditionally treated in connection with the Square of Opposition.

In general, Uzquiano (2017) defines quantification, and precisely its tool quantifiers, as "marks of generality". The most common examples of quantifiers in English include the determiners *all*, *each*, *some*, *many*, *most*, *no*, *a*, *few*, but also, e.g., *at most four*, *more than six*, etc. Quantifiers often occur with singular or plural nouns, in some cases combined with adjectives and relative clauses such as, e.g., *some good days*, *all students*, etc. Further, the phrases with quantifiers can be combined with predicates and form sentences such as *All students are happy*, or *Some students are female*. In natural languages, a noun, when it is clear from the context, can sometimes be left out in cases such as *All went* or *Few went*. Therefore, the quantifiers can occur as unary quantifiers, requiring adding only one thing, or as binary quantifiers which form a sentence from two parts. The latter can be formalized as *Q(A, B)* where *Q* is a quantifier and *A* and *B* the predicates. The binary quantifiers at best capture the grammatical form of English sentences.

As already mentioned, the study of quantification and precisely, binary quantifiers and their relations, including among others, contradictories and contraries, originates in Aristotle's *De Interpretatione* 6-7 and continues in *Prior Analytics*. Aristotle, however, focused on a restricted number of patterns which he referred to as syllogisms, including *Every S is P* (universal affirmative), *No S is P* (universal negative), *Some S is P* (particular affirmative), and *Some S is not P* (particular negative). Aristotle signified these logical forms as *A*, *E*, *I*, and *O* respectively. The logical relations which exist between them are based on the laws of logic and are illustrated in the Square of Opposition (Figure 1). At the heart of the Square of Opposition lay the LNC and the LEM.

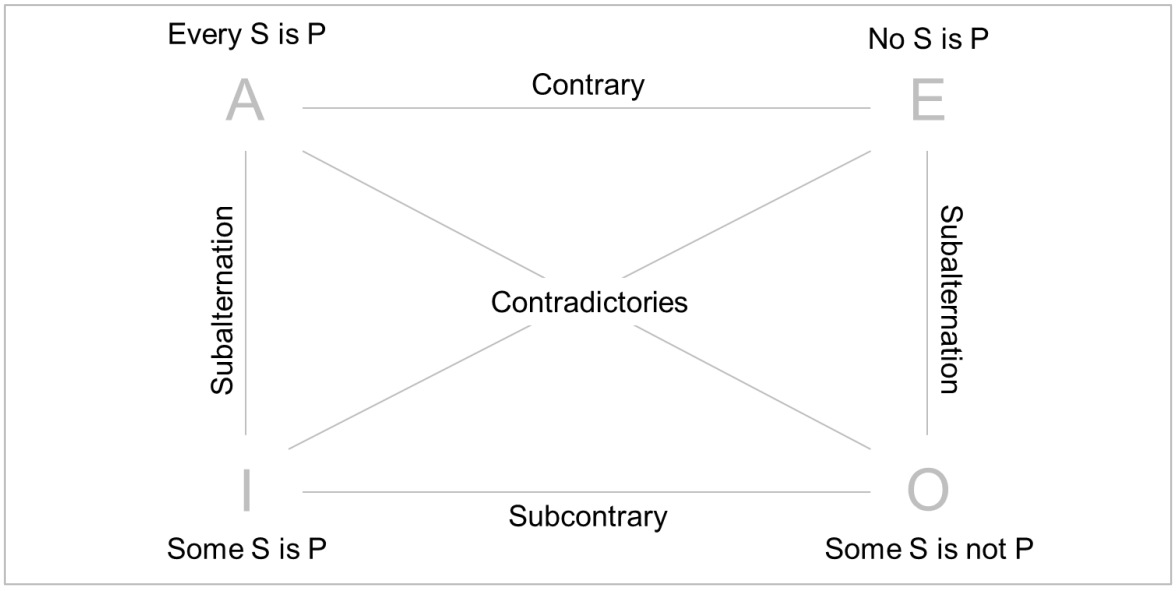

Figure 1: Square of Opposition.

One should note that in *De Interpretatione* 6-7, Aristotle only defines the relations between *A* and *O* (as contradictories), *E* and *I* (as contradictories) as well as *A* and *E* (as contraries). The correspondent text passage is as follows:

> "I call an affirmation and a negation contradictory opposites when what one signifies universally the other signifies not universally, e.g. every man is white – not every man is white, no man is white – some man is white. But I call the universal affirmation and the universal negation contrary opposites, e.g. every man is just – no man is just. So these cannot be true together, but their opposites my both be true with respect to the same thing, e.g. not every man is white – some man is white."

Other relations in the square are subsequently determined by a chain of implications, which is given e.g. in Parsons (2017).

Though the syllogistic logic of Aristotle is often referred to in literature, it is "eventually revealed itself as inadequate for the representation of mathematical argumentation" and was replaced by quantificational logic initiated by Boole's algebraic approach to logic and Frege's linguistic approach. For more elaboration on this, see Uzquiano (2014) and Parsons (2017).

3.1.4 Related Terms

Two related terms that should be mentioned in the context of contradiction discussion are those of tautology and paradox.

Two kinds of definitions of tautology exist. According to the first one, tautology can be defined as a statement which repeats in a verbal form something that has already been said previously (Apel 1976; Kondakov 1983; and Blackburn 2016). Apel (1976) refers to the latter observation as "saying the same". According to the second kind of definitions provided, e.g., by Kondakov (1983), Cook (2009), and Blackburn (2016), tautology is defined in terms of propositional or predicate logic and truth values as a compound statement which, independent of the truth value of its parts, is always true. Blackburn (2016) additionally states that "a tautology is thus valid, or true in all interpretations". A tautology can be illustrated by the following example (3.2):

(3.2) If Socrates is human, then Socrates is human. (Audi 2015)

When applying the second definition of tautology to the relation between tautology and contradiction, it can be observed that, while the truth value of contradictions is necessarily false,[14] tautologies are necessarily true in all interpretations (Lyons 1977: 787). A proposition that is the negation of a tautology is called a self-contradiction (Audi 2015). It is interesting to note that according to Mučnik (see Section 3.4.2), tautologies occurring in the text are psychologically perceived by humans as contradictions. The corpus-based study conducted in Chapter 5 provides an empirical evidence proving this claim.

[14] Please consider, this statement refers to the logical view on contradiction. From the dialectical point of view, there exist true contradictions. Moreover, it is to add, that there are also discussions whether there exist true logical contradictions (see for example Priest 1998; Parsons 1990).

Beside tautology, paradox is another term related to contradiction. Also for paradox, two views exist – a general and a logical one. According to the general view, paradox can be defined as something that is not commonly accepted, obviously senseless and illogical (Apel 1976; Kondakov 1983; Regenbogen/Meyer 1998). According to the second, logical view, paradox "is an argument that proceeds from apparently true premises, through apparently unobjectionable reasoning, to a patently false or even contradictory conclusion" (Cook 2009: 214). A similar definition is also provided in Regenbogen/Meyer (1998) and Blackburn (2016). According to Apel's (1976) logical definition of a paradox, it is a statement which is apparently both true and false at the same time.

A prominent example of a paradox is the liar paradox, which arises in case a sentence states its falsity, such as in (3.3). If this statement is true, then it is false, and vice versa, if the statement states that it is false, then it is true. In every case, the truth value assignment leads to a contradiction.

(3.3) This sentence is a lie.

In order to solve a paradox, Blackburn (2016) lists the following possibilities: a) to show that "there is a hidden flaw in the premises", or b) to demonstrate that there is an error in reasoning, or c) to disclose "that the apparently unacceptable conclusion can, in fact, be tolerated".

A variety of paradoxes are found in different fields, including biology, mathematics, physics, politics, etc. In logic, the paradoxes can be divided into two groups based on one side, on the aspect of self-reference (e.g., the liar paradox, Socratic paradox, barber paradox, or crocodile paradox) and vagueness on the other (e.g., the Sorites paradox, or paradox of the heap). Besides this distinction, logical paradoxes are further divided into logical and semantic paradoxes. Semantic paradoxes are represented by antinomy and aporia. An example of a semantic paradox that is antinomy is the liar paradox as previously illustrated in (3.3). Besides logical and semantic, also metaphysical and rhetorical (oxymoron) paradoxes exist.

3.2 Negation in Natural Languages

3.2.1 Typology of Negation

"All human systems of communication contain a representation of negation. No animal communication system includes negative utterances, and consequently, none possesses a means for assigning truth value, for lying, for irony, or for coping with false or contradictory statements." (Horn 1989).

In propositional logic, negation is considered in terms of truth value. That is, it is defined as an operator ¬¬ (*not*) that reverses the truth value of proposition p to proposition ¬*p* (*not p*),

which can be illustrated as a strict rule as in (3.4a). Another property of the logical negation is that it can cancel itself out, that is, to infer *p* from proposition *not-p* by negating it (the Law of Double Negation), and vice versa, without any changes in meaning. The formal representation of this rule is shown in (3.4b).

(3.4) a. If p is true, then not-p is not true

b. not(not-p) = p

Though natural language negation reflects the properties of a logical negation, it has, however, many more realization devices in its command, contributing to the expression of subtle nuances in meaning, being far from "simple and transparent" (Givón 1993). Consider the following example (3.5) taken from Givón (1993: 188):

(3.5) a. I am happy.

b. I am *not* happy.

c. I am *un*happy.

d. I am *not un*happy.

According to the logical rules (3.4a) and (3.4b), the examples in (3.5a and 3.5d) are synonymous as well as the examples in (3.5b) and (3.5c). However, from the speaker point of view, the examples (3.5b) and (3.5c), though both are negations of (3.5a), are not equivalent, incorporating slight differences in meaning, that is, expressing a different degree of happiness.

As already mentioned above, natural languages are characterized by a variety of devices in their command for expression of negation. Dahl (1979), Payne (1985), and Dryer (2005), for example, distinguish between three types of natural language negation, which with some variations in terminology, include negative particles, negative verbs, and morphological (or affixal) negation. In addition, Payne (1985) regards negative nouns as a distinct type of negation as well. Concerning what status should be given to double negative particles (e.g. *ne…pas* in French), different opinions exist. While Dryer (2005) and Dahl (1979, main text) propose to treat double negative particles as a distinct type, Payne (1985) and Dahl (1979, Appendix A) regard them as a variation of negative particles. In the present study, the double negative particles will be treated following the second view.

Negative particles (also called negative adverbs) are independent non-inflected words. In English, *not* and *n't* belong to the negative particles. Negative particles occur in about half of the world's languages (Dryer 2005) and are considered as "the most common type of standard negation" (Dahl 2010: 19).

From a linguistic perspective, some interesting aspects of negative particles are to be considered. One aspect is related to the position of negative particles within a sentence. The aspect of negation and the word order will be discussed in Section 3.2.2.

Another aspect encompasses the variants of the negative particle construction. Thus, e.g., the negation of an affirmative sentence by means of a negative particle can lead to a change in the verb's original form. Dahl (2010), with reference to Capell and Hinch (1970: 67), illustrates this case with the example of Maung, also Mawung, Mawng, and Gun-marung (Australian aboriginal language) (3.6), where the verb in the negative sentence is negated by the particle *marig* in combination with an irrealis suffix added to the verb:

(3.6) a) ŋi-udba-n
1sgA:3sgO-put-past:realis
I put it.

b) marig ŋi-udba-nji
NEG-1sgA:3sgO-put-past:irrealis
I did not put it.

Another variant of the negative particles is represented by double negative constructions such as the French ***ne**…**pas*** in e.g. *Je ne sais pas* (I don't know), which is a negation of *Je sais* (I know), and the Middle English ***ne**…**not***. This kind of negative particle which consists of two units is also referred to as discontinuous negation in de Swart (2010). As de Swart (2010) further points out, though the discontinuous negation consists of two units, it should be regarded as one negation cue which, in terms of logic, can be represented by the operator ¬. One should note that in spoken French, the preverbal *ne* is often dropped (de Swart 2010).

Besides French and Middle English, the use of double negative constructions could also be found in other Romance languages as well as in Celtic, Mayan, and West African languages (Dahl 2010; Dryer 2009).

Another interesting issue regarding the negative particle is the development of its double negative construction and the subsequent disappearance of its second element. This observation has been well described from the historical point of view in Jespersen (1917: 4):

> "The history of negative expressions in various languages makes us witness the following curious fluctuation: the original negative adverb is first weakened, then found insufficient and therefore strengthened, generally through some additional word, and this. in turn. may be felt as the negative proper and may then in course of time be subject to the same development as the original word."

Beginning with Dahl (1979), this process is referred to as ***Jespersen's cycle***. An analysis of Jespersen's cycle in the framework of the optimality-theoretic model is discussed in de Swart (2010). The development of negative particles in English and French is described in Horn (1989).

Beside negative particles, the negation in some languages is realized by means of negative verbs. There are two kinds of negative verbs, including higher negative verbs and auxiliary negative verbs. While the first type of negation is expressed as a combination of a verb with a sentential element, the second type involves a negative element which takes over the inflectional categories of finite verbs.

As Payne (1985) points out, only a few languages use negative verbs to express negation. That is, the use of higher negative verbs could be found in the Squamish language spoken in southwestern British Columbia, Canada (Kuipers 2015) and Malayo-Polynesian languages. Auxiliary negative verbs are used, e.g., in Estonian and Finnish languages. In Finnish, negation is realized by the verb *en* as in (3.7). For studies of negative verbs in Finno-Ugric languages, see also Mitchell (2006) and Kaiser (2006).

(3.7) a) Pekka lukee (Dahl 2010: 21)
Pekka-read-presence-3sg
Pekka is reading.

b) Pekka ei lue
Pekka-NEG-3sg-read
Pekka is not reading.

Finally, morphological negation (or affixal when speaking in terms of Zimmer (1964)) is mainly realized as an affix, commonly on a verb or an auxiliary. In Turkish, the standard negator is a suffix *-m-* in combination with a vowel. In English, morphological negation is realized by prefixes *in-*, *im-*, *il-*, *dis-*, *un-*, and postfixes *-less* and *-out* as in *without*. Dahl (2010: 14) notes that while the negation affix in Turkish is a kind of sentential negation, English prefixes should be regarded as lexical negation. Thus, Dahl (1979) clearly opposes morphological negation to syntactic negation, while Payne (1985) and Dryer (2005) regard morphological negation on the par with syntactic negation, such as in the case of negative particles.

In general, Dahl (1979) found that suffixal negation is used more frequently than prefixal negation, being in proportion of 1.75:1. Bybee (1985: 177), in contrast, found the slight prevalence of prefixal negation on the basis of a sample of 50 languages. Additionally, in some languages, such as Igbo (Green/Igwe 1963 cited in Dahl 2010), morphological negation is expressed by an affix in combination with a tone change, while in Liberian Mano (Becker-Donner 1965 cited in Dahl 1979) and Nigerian Mbembe (Barnwell 1969; Dahl 2010) through tone alone. Further ways of negation marking are realized by the tone and verbal reduplication as in African Eleme (Bond 2016) or through reduplication as in Caucasian Tabasaran (Khanmagomedov 1967).

The use of morphological negation is common to about a third of languages in the world. Bybee (1985) estimates its frequency to 30% and Dahl (1979) to 45%, while 33% account for negative affixes, according to estimates made by Dryer (2005).

Another typology of negation is proposed in Miestamo (2005a, 2005b, 2006). In comparison to the previously mentioned studies, Miestamo grounds his typology not on the nature of negative markers but on structural differences, or asymmetries, between negative and affirmative sentences. In this context, the researcher distinguishes between symmetric and asymmetric negation, characterizing the latter as "more complex than symmetric negation" Miestamo (2006: 345). He further divides asymmetric negation into "A/Fin" (asymmetry in the finiteness of verbal elements), "A/NonReal" (negative clauses are realized as non-realized, so-called irrealis), and "A/Cat" (negation involves changes in grammatical categories such as mood, person, tense, and aspect). That is, with respect to Miestamo's typology, the most negative particles can be regarded as symmetric negation. Negative verbs, in turn, are asymmetric and fall under the type A/Fin/NegVerb. Particularly worth mentioning is that according to Miestamo, some constructions can simultaneously involve asymmetries of different types. Regarding this, Dahl (1979: 12) argues that Miestamo "classifies asymmetries rather than negation constructions".

Typologies of negation with a focus on English have been proposed, among others, in Huddleston (1984); Givón (1993); Hidalgo Downing (2000); and Quirk, Greenbaum, and Svartvik (2010).

Thus, Huddleston (1984) distinguishes between clausal and subclausal negation such as in *She had had complete faith in no man* and subclausal negation as in *She had solved the problem in no time*, respectively. The main difference between the two types lies in the observation that subclausal negation "does not affect the polarity of the clause as whole" (Huddleston 1984: 423).

Similar to Huddleston (1984), Quirk et al. (2010) also propose a typology of negation based on a syntactic difference and distinguish between clause, local, and predication negation. Clause negation, which Quirk et al. (2010: 775) define as negation "through which the whole clause is syntactically treated as negative", includes negative particles such as *not*, *n't* and the pronouns *no one*, *nothing*, *nobody*, *never*, and *nowhere*. Furthermore, the researchers count to syntactic negation words which are negative in meaning but not in form. These include the adverbs and determiners *hardly*, *scarcely*, *seldom*, *rarely*, *few* and *little*, as well as to some extent, *only*. Also, verbs, adjectives, and prepositions with negative meaning such as *to deny*, *to forget*, etc., followed in particular by *any* and *ever*, are counted as syntactic negation. Local negation, in turn, is defined as that which negates "a word or a phrase without making the clause negative" such as in *She's a not unattractive woman* (Quirk et al.

2010: 791). Finally, predication negation, which is rarely used, however, occurs after certain auxiliaries which are “used with a different scope of negation than is normal for that auxiliary” (ibid., 797) such as in *They may not go swimming* (=They are allowed not to go swimming) with a pause before not.

Givón (1993), in turn, distinguishes between syntactic, morphological, and inherent negation. Syntactic negation in Givón (1993) is equivalent to the clause negation addressed in Quirk et al. (2010) and the sentence negation discussed in Klima (1964). Morphological negation, in turn, corresponds to the morphological negation as described in Dahl (1979, 2010), Payne (1985), and Dryer (2005). Finally, inherent negatives, which have not been addressed in Dahl (1979, 2010), Payne (1985), and Dryer (2005), are represented by words which have no overt negation marks. Strictly speaking, these are words which are positive in form but inherit negative meaning. Examples of inherent negation are the verbs *to fail*, *to absent*, *to lack*, *to forget*, *to deny*, *to refuse*, *to reject*, *to doubt* and the adjective *sad*, *bad* etc. Inherent negation or inherent negatives have been first discussed in Jespersen (1917) and also addressed in Quirk et al. (2010), Horn (1989), and Hidalgo Downing (2000). The process of identification of inherent negatives by a human from the theoretical and practical point of view have been discussed in Jespersen (1917: 43), Huttenlocher and Higgins (1971), Sherman (1973) and Clark, H. (1974), Clark H. and Clark E. (1977), and Tottie (1991: 7). Thus, Jespersen (1917: 43) regarded the identification of inherent negatives as a reversing process, such as in case of *to fail* and *to succeed*, where *to fail* is theoretically utilized as *not succeed*, and *succeed* as *not fail*. The same view on the processing of inherent negatives is shared in Clark (1974) as well as in Clark H. and Clark E. (1977). Horn (1989: 522), in turn, regards inherent negation as pragmatically complex as it relies on knowledge of earlier discourse.

The time needed by a human for processing of negatives and affirmatives, in general, as well as inherent negatives, was studied, e.g., in Sherman (1973) and Clark (1974). The results of experiments reported in Clark (1974) showed that true or false negatives (*isn’t present*) need more time to be processed than true or false affirmatives (*it’s present*). The same results have been obtained when processing the inherent negatives. Thus, e.g. the processing of the verb *to absent* followed the same pattern and duration of verification as *isn’t present* had.

By analogy with Givón (1993), Hidalgo Downing (2000) distinguishes between syntactic, morphological, and inherent negation, and he further groups these three types into explicit and implicit negation. According to this division, syntactic negation represents the explicit type of negation. Though the adjuncts *hardly*, *scarcely*, etc., convey a negative meaning without having any overt negation markings, the researcher regards them as explicit. In

turn, the scientist counts morphological negation and inherent negation to the implicit type of negation, which Hidalgo Downing also treats under the term lexical negation.

3.2.2 Negation and Word Order

The study of negation and its position in a sentence has been studied, among others, in Jespersen (1917), Greenberg (1966), Bartsch and Vennemann (1972), Lehmann (1974), Dahl (1979), and Dryer (1988). The main focus of the studies was thereby on the position of negation with respect to the verb. Traditionally, one distinguishes between a preverbal position (negation precedes the verb) and a postverbal position (negation follows the verb). The patterns of these two types were first described in Jespersen (1917).

To explain the choice of negation placement, most researchers built their theory on the position of negation in relation to the subject (S), object (O), and verb (V), and outgoing from the word order patterns of a certain language following Greenberg (1963). Thus, Lehmann (1974), e.g., argues that preverbal negation occurs in SVO languages, while postverbal negation is typical for SOV languages. Bartsch and Vennemann (1972), in turn, are of different opinion, claiming exactly the opposite of what Lehmann (1974) states.

The study of Dryer (1988), which was based on a sample of 345 languages, showed that in SOV languages, both preverbal and postverbal negations are common while the position of negation before the subject and object are not frequent. For SVO languages, in turn, the use of preverbal negation is characteristic. Also, languages with an initial verb, the so-called VSO and VOS languages, tend to position the negation before the verb.

In general, Jespersen (1924: 297 quoted in Horn 1989: 293) observed a strong tendency for languages to place the negation before there verb and saw a possible explanation for this finding in the intention of the speaker “to put the negative word or element as early as possible, so as to leave no doubt in the mind of the hearer as to the purport of what is said”. The tendency to place negation before verb was also observed in the study of Dryer (1988). From the sample of 345 languages, Dryer (1988) found that 70% of languages use a preverbal negation.

In case of the double negation marker, such e.g. French *ne…pas*, the units of the negation have different positions in the sentence, usually one preceding and one following the verb as in *Je ne sais pas* (=I don’t know).

3.2.3 Scope of Negation

Hidalgo Downing (2000: 45), following Quirk et al. (2010: 787-790), Givón (1993: 197-198), Huddleston (1984: 428-432), and Downing and Locke (2006: 25) define scope of negation as “the semantic influence that the negative item exercises over the constituents of the clause where it appears, or the semantic domain on which negation applies”. In other words,

the discussion on the scope of negation aims at identifying the means which indicate "what it is that is being negated semantically" (Huddleston 1984: 428).

It is commonly considered that all constituents that follow the negative word in the main clause fall under the scope of negation such as in (3.8a and 3.8b).

(3.8) a. Ed hasn't read the book. (Huddleston 1984: 428)

b. Ed wasn't there, though he had said he would be.

Following the rule, in (3.8a), it is negated that *Ed has read the book*, which can be presented as not + {Ed has read the book}. In (3.8b), however, it is only negated that *Ed was there*, leaving the *though*-clause outside the scope of negation.

Certain modal auxiliaries – in all or only some of their senses – in the normal case, fall outside the scope of negation (Huddleston 1984). Thus, e.g., the verbal modals *must* and *ought* always fall outside the scope of negation, independent of their senses. That is, the sentence *You mustn't do that* in (3.9) has the interpretation of (3.9a) rather than of (3.9b).

(3.9) You mustn't do that. (Huddleston 1984: 430)
a. =You are required not to do this.
b. =You are not required to do this.

The verb *may*, in turn, falls outside the negation scope only in its epistemic sense (see Section 3.3.2 on the definition of modality). In contrast, *may* in its deontic sense falls within the scope of negation. Following this, the example given in (3.10) has to be interpreted rather as (3.10b) than as (3.10a).

(3.10) You may not have any more. (Huddleston 1984: 430)
a. =You are permitted not to have any more.
b. =You are not permitted to have any more.

A problematic issue in identifying the scope of negation is posed by the simultaneous occurrence of negation and quantification. The studies of the use and behavior of negation and quantification include, among others, Kahrel (1996), Haspelmath (2005), and Penka (2011). Kahrel (1996) and Haspelmath (2005), e.g., are among the first to reveal variations in usage of quantification and negation and propose a typology on the basis of their findings. Regarding the identification of the scope of negation, the simultaneous occurrence of quantification and negation, in most cases, leads to ambiguous interpretations of the sentence meaning as illustrated in (3.11).

(3.11) They hadn't processed one of the applications. (Huddleston 1984: 430)
a. =One of the applications hadn't been processed.
b. =Not one of the applications had been processed.

Aside from the prosodic properties of the sentence, the latter can be ambiguously interpreted as (3.11a) and (3.11b). In the first interpretation, the quantification expressed by *one*

is outside the scope of negation while in the second interpretation, the quantification is within the scope of negation.

Also, the occurrence of adverbs such as *very*, *always*, *often*, *sometimes*, *ever*, *clearly*, etc., can indicate the scope of negation. That is, very e.g. falls inside the negation scope only if the word which it modifies falls within the scope of negation, as illustrated in (3.12a and 3.12b), while *ever* always falls within the scope. *Always*, *often*, and *sometimes* behave like *all*, *many*, and *some*, respectively, with exception of *always*, which falls inside the scope of negation only when it follows a negative expression, and of *sometimes* for which it is less usual to fall inside the negation scope than for *some* (Huddleston 1984: 431).

(3.12) a. He very wisely didn't accept. (Outside the scope of negation)

b. He didn't behave very wisely. (Inside the scope of negation)

In addition to the above-presented cases, the negative scope can be indicated by means of a contrastive stress, narrowing the focus to a particular constituent in the sentence and leaving the rest of the clause presupposed (Quirk et al. 2010: 789; Givón 1993: 197) as illustrated in (3.13).

(3.13) a. John didn't hit Bill. (Hidalgo Downing 2000: 45)

b. *John* didn't hit Bill.

c. John didn't hit *Bill*.

Meanwhile, (3.13a) has a neutral focus so that the general rule for identification of the scope of negation is applied. In (3.13b), it is the subject which is in focus indicating that some person hit Bill, but this person is not John. In turn, in (3.13c), the object is in the focus, expressing that John hit some person, but this person is not Bill.

Further, Huddleston (1984), Givón (1993), and Downing and Locke (2006) emphasize a particular behavior of adjuncts occurring in the sentence with negation, forcing the former to attract the focus of negation while the rest of the sentence remains presupposed. That is, in (3.14), only the adjunct *as fast as she could* falls under the scope of negation, and not the verb *run*, leaving the presupposition *she ran* without changes.

(3.14) She didn't run as fast as she could. (Hidalgo Downing 2000: 46)

Finally, as Downing/Locke (2006) point out, in the case of two negative words occurring together in the same sentence, each negative item has its own scope as in the example (3.15).

(3.15) You can't NOT go. (Downing/Locke 2006: 25)

3.2.4 Double and Multiple Negation

The discussion of multiple negation – a situation when two or more negative elements are used in a sentence – is traditionally grounded on the distinction between the principles of duplex negatio affirmat (DNA) and duplex negatio negat (DNN). Horn (2010) refers to these principles also as logical double negation and hypernegation, respectively (see Figure 2).

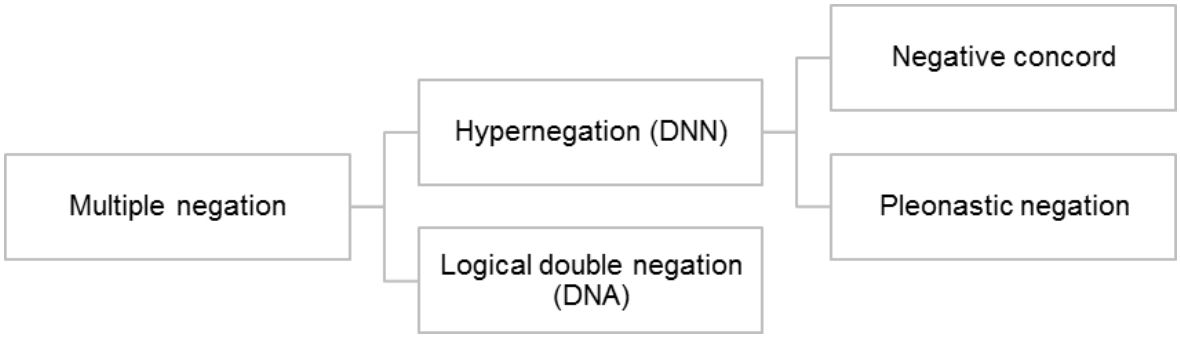

Figure 2: Multiple negation in English and other languages.

The DNA principle states that two negative elements are resolved, or canceled out, by each other, which results in the affirmation as shown in (3.16). The rhetorical figure litotes is grounded on the DNA principle. Two negative elements occurring together which express affirmation is typical for languages such as German and Latin. Also for standard English, the DNA principle is commonly the case. Bishop Lowth (1762: 126 cited in Horn 2010: 111) formulated this in his *A Short Introduction to English Grammar with Critical Notes*, stating "two negatives in English destroy one another, or are equivalent to an affirmative".

(3.16) Ich habe nicht nichts gegessen (German)
=Ich habe etwas gegessen
I-have-not-eaten-nothing
=I-have-eaten-something

The DNN principle, in turn, is built on the idea that two (or more) negative elements, when combined, express a single negation, achieving its intensification by this means. The DNN principle is found in the Romance and Slavic languages, as well as in non-standard English, and can be illustrated as in (3.17) for Russian and (3.18a and 3.18b from Labov 1972: 773) for Non-standard and Standard English.

(3.17) Ya nichego ne videl (Russian)
I-nothing-not-saw
=I didn't see anything (Standard English)

(3.18) a. It ain't no cat can't get in no coop (Non-Standard English)
b. = There isn't any cat that cannot get into any (pigeon) coop (Standard English)

While (3.18a) can occur in some dialect of English, it is in most cases, not accepted in formal situations and (formal) written language. Regarding Standard English, the correct formulation of (3.18a) would be, therefore, (3.18b). One should, however, note that the DNN principle can also be accepted in some stylistically motivated cases such as in (3.19).

(3.19) I can't get no satisfaction (from the song *(I Can't Get No) Satisfaction* performed by The Rolling Stones)

In general, languages for which DNN is applied are said to encompass the phenomenon of negative concord (NC) as well as some other relations across clause boundaries, which Horn (2010) treats under the term of pleonastic negation.

NC refers to cases where "multiple occurrences of morphologically negative constituents express a single semantic negation" (Labov 1972) as already shown above in (3.17). According to Haspelmath (1997), NC is more common among languages and, besides in non-standard English, is found in the Slavic (e.g. Russian) and the Romance languages (e.g. Spanish, Italian), also in Afrikaans, Japanese, and some dialects of German. In turn, languages without NC (DNA languages), such as Standard English, have Negative Polarity Items (NPIs) at their disposal.

NPIs, as defined, e.g., in von Bergen and von Bergen (1993) and van der Wouden (1994), are lexical items which are restricted to occur only in negative contexts, or negative sentences. NPIs in English include, in addition to some and any, also the auxiliary verb *to need* as well as *anything*, *much*, *many*, *ever*, *far*, *long*, *to budge*, *to lift a finger*, etc. The negative contexts, where NPIs can occur, according to Bolinger (1960), Klima (1964: 311-315), Borkin (1971), Karttunen (1971), Reinhart (1976), Horn (1978), Ladusaw (1980), and Linebarger (1981: 68-77, 1987: 362-381), include the following cases (so-called triggers of negative contexts), which are summarized with examples given in von Bergen and von Bergen (1993: 28-31):

- Explicit negation (Nobody expects him to write *any* more novels);
- Interrogative clause (Does she like him at *all*?)
- Conditional clause with *if* (If you *give a damn* about the whales, you'll contribute);
- Conditional clause with *unless*, *so far as*, *in case* (He will not be able to pay his debts unless *anything* turns up);
- Comparative clause (This outcome is more serious than John *ever* imagined);
- Subordinate clause with *before* and *long after* (She persisted long after she had *any* hope *at all* of succeeding);
- Relative clause after a non-specified nominal clause (That's all he *ever* thinks about);
- Relative clause with a superlative or after *first*, *last* (He was the first scientist who *ever* published *anything* on this subject);
- *too* + adjective/adverb + infinitive (He was too frightened *to say a word* against it);
- Complements of adversative, including morphological negatives, inherent negatives, affective predicates, such as *annoyed*, *surprised*, *amazed*, adjectives

expressing difficulty degree such as *difficult*, *hard*, *tough* (We're surprised that *anyone* bought *anything at all*);

- Negative sentence adverbs *rarely*, *seldom*, *hardly ever*, *scarcely* (John *hardly ever* lifts a finger to help us);
- Occurrence of *only*, *exactly*, *at most* (Only John has *any* interest in playing soccer);
- Negative prepositions *without*, *against* (He was against doing *anything* like that);
- *It's been* + time expression + *since* (It's been a week since I bought *any* book);
- Emphasis (No one lifted a finger to help us – John DID *lift a finger*).

One should note that the provided list of triggers of negatives contexts is in no case complete (von Bergen/von Bergen 1993: 32).

All the above-provided examples express a negative idea (=negative context), which is a necessary condition for the occurrence of NPIs. In practice, however, the negative context is not always easy to identify. Horn (1978: 152) refers to this as "buried negatives": "It should be clear that the trigger properties of even the least obviously negative elements result directly from the buried negatives they…contain". Another problem arising in the discussion of negative contexts is the definition of a negative context, that is when a context can be considered to be negative. Von Bergen and von Bergen (1993: 34), for example, argue that defining negative contexts as those which allow the occurrence of uncontroversial standard NPIs, such as any and ever, represents a sufficient criterion for the identification of negative contexts.

To continue the discussion of NC, the researchers distinguish between three types of NC, including some slight differences in terminology and definitions, and these are strict NC (also negative doubling, not to confuse with double negation), non-strict NC, and negative spread (de Swart 2010: 44).

Van der Wouden and Zwarts (1993: 202), following den Besten (1986), define doubling and negative spread as follows:

- Negative doubling is obligatorily present in all sentences containing a negative expression;
- Negative spread is observed if "the negative feature is 'spread' or distributed over any number of indefinite expressions within its scope.

Negative doubling is typical for French, Czech, Polish, Greek, Romanian, Hungarian, and Japanese; an example is shown in (3.20), which is taken from van der Wouden and Zwarts (1993: 202). Negative spread, in turn, is found in e.g. non-standard English and is illustrated in the example above (3.18a).

(3.20) Je *n*'ai vu *personne*. (Standard French)

I-not have-seen-anobody
=I haven't seen anybody

Giannakidou (2017: 9), in turn, distinguishes between strict and non-strict NC, which she defines as follows:

- Strict NC: The n-word[15] always requires the presence of the negative marker, regardless of position in the sentence;
- Non-strict NC: The n-word can appear without the negative marker in preverbal position or when construed with another preverbal n-word.

Strict NC corresponds to the negative doubling as defined in van der Wouden and Zwarts (1993). As already mentioned above, strict NC is found in Slavic languages, Romanian, Hungarian, as well as in several East Asian languages (e.g., Japanese, Korean), while most Romance languages, e.g., Spanish, Italian, and European Portuguese are examples of non-strict NC (de Swart 2010; Giannakidou 2017). In these languages, a postverbal n-word requires the presence of a preverbal marker of negation as shown in the example of (3.21a). In case the n-word is in a preverbal position, the negation marker is not used (3.21b).

(3.21) a. *Non* ha telefonato *nesssuno* (Italian; Giannakidou 2017: 9-10)
NEG has-called-NEG.body
=Nobody called.

b. *Nessuno* ha telefonato a nessuno.
NEG.body-has called to-NEG.body
=Nobody has called anybody.

Another realization of the DNN beside NC is represented by pleonastic negation, which Jespersen (1917: 75) defines as:

> "A negative is placed in a clause dependent on a verb of negative import like "deny, forbid, hinder, doubt". The clause is treated as an independent sentence, and the negative is expressed as if there had been no main sentence of the particular type."

As Horn (2010) points out, pleonastic negation is observed in case negative markers follow the comparatives, a clause with *before*, or verbs of fearing in languages such as, e.g., French, Russian, and Yiddish. Examples of pleonastic negation for English are provided in (3.22a, 3.22b, and 3.22c). It should be noted that pleonastic negation in some studies is also termed paratactic (Jespersen 1917), sympathetic (Smyth 1920/1974), and abusive (Vendryès 1950) negation.

(3.22) a. Well, really, how can I *keep from not* worrying? (Horn 2010: 125)

[15] The term n-word originates in Laka (1990) and refers to negative words which begin with *n-*, such as Italian *nessuno*, Spanish *nadie* and Portuguese *ningu* and appear in NC structures (Giannakidou 2017: 6-7).

b. I can't *keep from not* thinking about the impending doom of it all.

c. We sure *miss not* seeing you every day, Bob.

Whereas most researchers primarily focus on typology construction, Horn (2010) is also interested in the general considerations related to double and multiple negations posing, among others, the following questions:

- Why use multiple negative elements to express negation if, from a semantic point of view, one would be sufficient? What is the motivation?
- Why would a speaker make an effort of using two negations to express an affirmation?
- What do multiple negations in a context of DNA affirm?

Horn (2010: 113) puts in remembrance that Grammarians "condemn the use of logical double negation as a marginal, superfluous, and suspiciously Latinae phenomenon" and also mentions, among others, the names of Lowth (1762), Orwell (1946/1961), and Tesniére (1959/2015). This negative view on the DNN opinion, in a sarcastic manner, is well expressed in Orwell's (1946: 357, 365 cited in Horn 2010: 113) quote:

> "Banal statements are given in appearance of profundity by means of the *not unformation*…It should be possible to laugh the not un formation out of existence…One can cure oneself of the not un formation by memorizing this sentence: *A not unblack dog was chasing a not unsmall rabbit across a not ungreen field*."

As possible motives for use of multiple negation in order to construe affirmation, Horn (1991, 2010) names quality, politeness or diffidence, weight or impressiveness of style, absence of corresponding positive, parallelism of structure, as well as minimization of processing, in contexts of direct rebuttal or contradiction. For more philosophical elaborations on the DNN as well as DNA, see e.g. Horn (2010).

3.3 Problematic Issues about Contradiction

3.3.1 Contradiction and Presupposition

As shown in Section 3.1.1, at the core of the semantic version of the LNC and, as a consequence, of contradiction lies the assignment of truth values. The latter, however, meets with obstacles, especially when dealing with the notions of presupposition and modality. The relation between modality and contradiction is the topic of Section 3.3.2.

Traditionally, presupposition is understood as a condition that must be fulfilled in order for the sentence to be judged true or false. Which truth value is the case for a sentence arises when the sentence is uttered in context. The idea of using the term presupposition originated in Frege (1892/2011) and was developed as a requirement for names to have a meaning (Vater 2005: 31). Consider the following example (3.23):

(3.23) a. Putin is the present president of the Russian Federation.

b. Putin is not the present president of the Russian Federation.

Uttering the sentences (3.23a) and (3.23b) in a context, the truth of the sentence can be determined under the condition that there is a person named *Putin*.[16] In other words, the truth of the sentence presupposes the existence of the referent of Putin at present. This condition must be fulfilled both for affirmation (3.23a) and negation (3.23b) in order that they can be assigned the truth value. If the referent of Putin, however, is not given in the context, it cannot be decided whether the sentence is true or false. This problem has first been introduced by Russell (1905/1988) and traditionally exemplified by (3.24)

(3.24) a. The present king of France is bald.

b. The present king of France is not bald.

Again, the truth of the sentence in (3.24a) presupposes that there is a king of France at present. That is, if France is a monarchy, then (3.24a) is either true or false. However, since France today is a republic with no king, it is not clear which truth value (3.24a) should be assigned to. The same situation is the case regarding negation (3.24b).

Philosopher and logicians proposed different approaches to deal with such kind of cases which are referred to as the Russellian problem. Thus, Russell (1905/1988), e.g., proposed always to regard sentences with false presupposition as false. Strawson (1950), in turn, criticized Russell's solution, claiming that it contradicts the laws of logic, and proposed treating such kind of sentences without assigning any truth value, regarding them as sentences with a truth gap (this term, however, was first introduced by Quine (1959)). The discussion as to which approach should be preferred, we will leave unanswered for now because truth values will not be considered for the development of a CD system.

Beginning with Stalnaker (1972, 1973, 1974, 1998), besides the truth-based notion of presupposition discussed, a pragmatic notion of presupposition has been established in theory. In contrast to the truth-based notion which focuses on what words and sentences presuppose, the pragmatic presupposition is understood through the prism of language usage, referring to the knowledge that people require as a precondition while speaking. As defined in Beaver and Geurts (2014), following Stalnaker, a pragmatic presupposition of a sentence "is a condition that a speaker would normally expect to hold in the common ground between discourse participants when that sentence is uttered". From that definition, it can be inferred

[16] Note, that presuppositions are not cancellable under negation. This ability of presupposition to remain under negation is considered its particularity and distinguishes presupposition from entailment and implicature (Section 5.2.2).

that presuppositions are not being expressed explicitly by the speaker (but triggered by certain lexemes and grammatical constructions) and thus, do not belong to the propositional content of the utterance. Pragmatic presupposition and contradiction will also be the focus of Section 5.2.2. The pragmatic notion of presupposition will only be partially taken into account for the purpose of the present study, that is, conceptually considered but not practically implemented.

3.3.2 Contradiction and Modality

Each sentence either asserts or negates the existence of a certain situation. However, in some cases, a situation can also be described as coincidentally or necessarily present or absent, and not merely as present or absent. Moreover, some situations or actions are described as allowed or not allowed, or as necessary. Relating to the future, situations can be presented either as possible and necessary, or as impossible. Finally, our knowledge about some action or situation can be judged proved or not, certain or not, etc. Depending on whether propositions contain such kinds of judgments, the former is divided into assertoric and modal.

The study of modal propositions was initiated by Aristotle. In Aristotelian logic, assertoric and modal propositions are termed as assertoric and problematic, respectively. In general, assertoric propositions as in (3.25a) can be defined as those which either assert or negate some information. In contrast, modal propositions as in (3.25b) neither strictly assert or negate some information but contain some evaluation or judgment of this information.

(3.25) a. The president of the U.S.A. John F. Kennedy *is killed*.

b. The president of the U.S.A. John F. Kennedy *must be killed*.

The expression of evaluation and judgment of information in a proposition, including necessity, permission, and probability, as well as their negation, is referred to in logic and linguistics as modality.

In general, logic and linguistic theory distinguish between the alethic (“necessary”, “possible”, “impossible”), deontic (“obligatory”, “permitted”), and the epistemic (“know”, “believe”, “imagine”) kinds of modality. Some researchers additionally distinguish between the temporal (“always”, “sometimes”, “never”) and the axeologic (“good”, “bad”) modalities. The terms are derived from modal logic in linguistics and are used, e.g., to study the behavior of modal verbs (Crystal 2008: 19).

The first type of modality – alethic modality as defined in Lyons (1977: 788-791), Crystal (2008: 19), and Palmer (1986: 10-11) indicates the logical necessity, possibility or impossibility of the proposition expressed by the speaker’s utterance shown in the examples (3.26) and (3.28b).

(3.26) A triangle must have three sides (Crystal 2008: 19)
=It is impossible for a triangle not to have three sides.

The second type, the epistemic modality, refers to the expression of certainty or the evidence that a speaker has for the proposition of his utterance (Crystal 2008: 171; Chung/Timberlake 1985: 242; Bybee 1985: 165-166; Palmer 1986: 10-11, 51) as shown in the examples (3.27) and (3.28c). Epistemic modality can be expressed in terms of "know", "believe", "doubt", "think" "imagine", etc. (Lyons 1977: 794). There are two kinds of epistemic modality, which are the evidentiality and judgment modalities. While evidentiality indicates the speaker's assessment of the evidence for his statement (Bybee 1985: 184; Palmer 1986: 66-67; Givón 1984: 307-308), the judgment modality refers to the speaker's strength of inference or degree of confidence (Palmer 1986: 53, 57-58, 64).

(3.27) He knows that Edinburgh is the capital of Scotland. (Lyons 1977: 794)

Finally, the deontic modality, as illustrated in (3.28a), expresses the logic of obligation, permission, or the modal desirability of the proposition expressed by the speaker's utterance (Lyons 1977: 823; Crystal 2008: 136; Palmer 1986: 10-11, 15, 96-97, 115). Deontic modality can be expressed in terms of "obligatory", "permit", "prohibit", "exemption". One distinguishes between three kinds of deontic modality, which include the commissive modality (indicating the speaker's commitment to do something like a promise or a threat), the directive modality (referring to the expression of commands and requests) and the volitive modality (which indicates the speaker's attitude of hope, wish, desire, or fear).

(3.28) The car must be ready (Crystal 2008: 136)
a. =It is obligatory that the car be ready (deontic modality)
b. =It is metaphysically necessary for the car to be ready (alethic modality)
c. =It follows from what is known that the car is ready (epistemic modality)

As already mentioned above, the difficulty with the treatment of modalized propositions consists in their inassertability. The modalized propositions express the modal attitude to the proposition they are applied to and are non-truth-functional. That is, it is impossible to express modal expressions in terms of two-valued logic. That is why the apparatus of traditional and classical logic fails to deal with modality, giving rise to non-classical logic. Thus, e.g., modal propositions can be treated within the three-valued logic proposed by Łukasiewicz (1920/1970).

A solution for the treatment of contradictions under epistemic modals was first proposed in Yalcin (2007). Yalcin (2007) claimed that the sentences of the form $\varphi \wedge E \neg\varphi$, where E is an epistemic modal, though unassertable, must be contradictory. To validate this idea, he developed the notion of informational consequence for validating this idea. Consider the following example taken from Yalcin (2007: 1):

(3.29) a. It is raining, and it might not be raining.

b. It is raining, and possibly, it is not raining.
c. It is not raining, and it might be raining.
d. It is not raining, and possibly, it is raining.

Yalcin (2007) claims that it is obvious, that the sentences in (3.29) sound odd, or "pragmatically defective". They are not assertable but seem to be consistent as well. Based on the semantics of the epistemic modals embedded in the sentences, it can be said that they both entail the sentences in (3.30). The sentences in (3.29) are odd, therefore, those in (3.30) are also odd.

(3.30) a. It's raining, but I don't believe that it is raining.

b. It's raining, but I believe it is not raining.

The traditional way to evaluate sentences of this kind in (3.30) would be to interpret them as Moore-paradoxical. That is, to deny that they are contradictions in any semantic sense, explaining this by their intelligibility (see Moore (1993) for more information). Yalcin (2007), in turn, proposes to treat such cases as contradictions or as epistemic contradictions.

Yalcin (2007) assumed that Moore's paradox and epistemic contradictions show different behaviors if they are put in some certain environment. Precisely, the researcher claimed that, while Moore's paradoxical sentences will not preserve their oddness under some certain environment, the epistemic contradictions will. This idea is referred to as Yalcin's puzzle.

To validate this idea, Yalcin (2007) embedded the sentences in (3.29) in the constructions *suppose* and *if* as in (3.31).

(3.31) a. Suppose it's raining and it might not be raining.

b. If it is raining and it might not be raining, then…

The imperative (3.31a) and the conditional (3.31b) sound unacceptable as we cannot coherently suppose that it is raining and that it might not be raining. Either way, it is unacceptable to say (3.31b). Therefore, according to Yalcin, here we observe an epistemic contradiction as the sentences remain odd, even when embedded under *if* and *suppose* construction. In the case of Moore's paradox, if we suppose that *it might not be raining* in sense of *I don't believe that it is raining*, or *I believe that it is not raining*, the sentences in (3.29) would be acceptable.

Schulz (2010) argues that Yalcin's approach is inadequate for justifying the existence of a contradiction construed by epistemic modals and proposes an alternative way to deal with them, which due to irrelevance in the present study, will not be further addressed here.

Besides in Yalcin (2007), the relation between modality and contradiction has been addressed in Svintsov (1979: 193) and in de Marneffe et al. (2008). Thus, Svintsov (1979)

claimed that sentences which include incompatible modality qualifiers such as *doubtless* and *beyond debate* on the one side and, e.g., *likely* and *probably* on the other, can be regarded as contradictory. De Marneffe et al. (2008), in turn, regarded the sentence pairs of the opposite modalities of possible and not_possible, of actual and not_actual, and of necessary and not_necessary as contradictory. According to the authors, the modalities were identified according to predefined modality markers such as *may*, *can* and *maybe*.

3.3.3 Contradiction and Vagueness

It is traditionally considered that the truth value of a contradiction is necessarily false. However, since the 80s of the last century, there have been discussions on the possibility of true contradictions existing (see e.g. Priest 1998, 1999; Parsons 1990). Contradictions that contain vague predicates enjoy a particular status in this discussion (see e.g. Fine 1975; Beall/Colyvan 2001; Ripley 2009; Cobreros et al. 2010; Alxatib et al. 2013). Before proceeding with this discussion, the definitions of the terms vague and vagueness will first be given.

As it can be observed when using a language in real life, a variety of lexemes do not have clear boundaries in their meaning. This, in turn, allows adapting them flexibly to a given context of utterance (henceforth CoU). That is, e.g., when referring to a child as a *baby*, the meaning of the word depends on gradable criteria such as the age and developmental stage. A child considered a baby by one person is not necessarily considered a baby by another. The denotation of the term *baby* has, in this case, flexible boundaries and can be freely adapted to a given CoU.

In general, Löbner (2013: 47) defines lexical vagueness as follows: “A lexical meaning is vague if it allows for flexible adaptation to the given CoU”. This is the reason why vague terms are often used in a deliberate rhetorical strategy to avoid dealing with an issue or directly responding to a question. Vagueness can be observed by using concepts with properties varying on a continuous scale such as, e.g., in the example of the term *baby* provided above or in case of color terms (yellow, green, red, etc.) where a human perceives “a continuum with fuzzy transitions” (Löbner 2013: 47). According to Löbner (ibid.), all gradable adjectives which include comparative and superlative forms are vague terms.

It is also noted that the term vagueness often occurs in combination with polysemy. That is, the term light has different meanings, referring to the degree of difficulty, weight, etc. These are two different scales, and for each scale, the word *light* can be flexibly interpreted.

The impact of vague language on the status of contradictions has been addressed in Ripley (2011), Sauerland (2011), Cobreros et al. (2010), Alxatib et al. (2013), and Snider (2015). Consider the following example taken from Alxatib et al. (2013: 620):

(3.32) John is tall and is not tall.

Presuming that the two statements refer to the same point of time and to the same person, the statements are considered, at first sight, to satisfy all the conditions for being evaluated as a contradiction as stated in Aristotle's ontological definition (Section 3.1.1). One statement is true, while the other is false. However, when taking the vague meaning of the predicate *tall* into consideration, it can be concluded that it is not clear whether *tall* is true or false for John. Regarding the scale of height with the two polarities *tall* and *not tall*, where is the starting point at which someone can be characterized as tall? It is obvious that the same height can be sufficient for playing football (=*tall*) but not sufficient for playing basketball (=*not tall*).

Ripley (2011), Alxatib et al. (2013), and Snider (2015) refer to cases such as in (3.32), which contain vague predicates, as contradictions at the border, or borderline contradictions, while Ripley was the first to propose the term.

A borderline contradiction is observed when two statements satisfy the conditions of contradiction as defined in Section 3.1.1 but have vague predicates. Formally, a borderline contradiction can be represented as $Fa \wedge \neg\neg Fa$, where *F* is a vague predicate and *a* is a borderline case of the predicate.

Besides the example discussed above, a borderline contradiction can also arise as a result of the lexical usage referring to a human dynamic (emotional) experience, such as, e.g., a pair of opposites, *love–hate* in the example below (3.48) in Section 3.5.2. Simultaneously by stating *I love you* and *I hate you* (which is theoretically assumed to mean *I don't love you*) toward the same person, this is not necessarily to be interpreted as a contradiction. In case of the I love you referring to a partner's patience and the *I hate you* (= I don't love you) that refers to the partner's laziness, the words *love* and *hate* are each used in a different sense. That is, it can be concluded that a borderline contradiction is observed here.

The study of speaker intuition behind dealing with the contradictions containing vague predicates was studied, among others, in Bonini et al. (1999) and Ripley (2011). The results of the experiments reported in Ripley (2011) show that both sides of borderline contradictions are usually accepted, that is, they are both perceived as true by ordinary speakers. Further, Ripley (2011) showed that in some borderline contradictions, one is more acceptable than the other. Thus, e.g. if told that 5'11'' is the height of a Western man, people will more likely accept the contradiction in (3.33a) than in (3.33b) (Alxatib et al. 2013: 622):

(3.33) a. A 5'11'' tall man is and isn't tall.

b. A 6'4'' tall man is and isn't tall.

Regarding the framework for justifying the existence of true contradictions, the case *a* is a borderline case and *F* is a vague predicate, then the conjunction of *a is F* and *a is not F* is

true as well as it is true in the case that *a is neither F nor not F*, different approaches have been proposed. Ripley (2009) justifies the truth of borderline contradictions based on the theory of vague language within the framework of the paraconsistent logic. Fine (1975), in turn, rejects the truth of borderline contradictions claiming that they are false even when containing vague predicates. In fuzzy logic, a golden middle view has been developed on the truth value of borderline contradictions in which they are considered partially true (Smith 2008). Further, Cobreros et al. (2010) propose a pragmatic account for dealing with the acceptance of borderline contradictions, based on speaker's evaluation of borderline cases operating with notions of strict, classical, and tolerant truth. Alxatib et al. (2013: 620), in turn, point out that the pragmatic account, for some cases, is not applicable as the latter can be evaluated both as strict and non-strict; they therefore propose a semantic approach for the analysis of borderline cases of contradictions, assuming "a) a multivalued logic, and b) a semantics of conjunction that is not fully truth-conditional, but contains a modal component".

From the perspective of the construction of artificial intelligence systems, borderline contradictions represent a challenging task. To deal with vagueness in the present study, it seems reasonable for us to introduce the notion of potential contradiction (Section 7.1).

3.3.4 The Concept of Fake Contradiction

According to Aristotle's ontological definition of contradiction, which is prior to the LNC, "It is impossible that the same thing can at the same time both belong and not belong to the same object and in the same respect, and all other specifications that might be made, let them be added to meet local objections" (*Metaphysics* IV 3 1005b19–23). In other words, parts of contradiction – assertion and its negation – must refer to the same subject at the same time and in the same respect. That is, the two statements must be equivalent, with the exception of one thing: what is asserted in one statement must be negated in the second. If these conditions are not satisfied, there is no contradiction to be observed.

Consider the following example (3.34) which is a quote by the Russian writer Chekhov, cited in Ivin (2015).

(3.34) In my childhood, I had no childhood.

Despite the author simultaneously stating that he had and didn't have a childhood, the reader or hearer, especially those familiar with Chekhov's biography, will in most cases, agree that there is no contradiction here. This kind of expression represents a kind of "play of rhetoric and aphorism" (Ivin 2015: 70). There is no actual contradiction here as there is no assertion and negation of the same thing in the same sense and in the same time reference.

Brought up in a poor family with many children, Chekhov, in his childhood, had a lot of life's challenges to cope with. Expressing this by using the word *childhood* twice, the writer in the first case, refers to a period of life that every human has (a prototypical sense of the word *childhood*) while in the second case, he refers to the experience which is commonly associated with this period of life. Thus, as the word *childhood* is used, each in a different sense, no contradiction is present in the quote since some obligatory contradiction conditions (primarily in respect to the ontological view on contradiction, see Section 3.1.1) are not satisfied. The function of the negation operator no as a contradiction-building element is canceled here. But also with respect to the semantic view on contradiction (Section 3.1.1), stating that *I had a childhood* and immediately negating it does not constitute a contradiction as both parts of the sentence are assumed to be true. Thus, the LNC has not been violated here.

In the literature (Svintsov 1979 and Ivin 2015 to name only a few), such cases where a contradiction can be resolved by addressing the context, conducting word disambiguation, and fixing the references are synonymously referred to in English as fake, seeming, or apart contradictions. In the present study, the term fake contradiction will be used.

As mentioned, fake contradictions can be recognized in context by fixing the references, as e.g., in the conjunction *Socrates is a man* and *Socrates is not a man*, where *Socrates* in the first case, refers to the Greek philosopher and in the second case, to another person. If the context is not given, and in case of a lack of information, a recognition of fake contradictions can be problematic.

Similar to contradictions at the border, fake contradictions represent a problem case in the development of an efficient natural language system for CD. It is a challenging task as a deeper interpretation is required for the meaning of text and context.

3.4 Classification of Contradictions

3.4.1 Classification of Svintsov

There exists only basic research on the classification of contradictions that naturally occurs in textual data, which was mainly conducted in the 80s. It is, however, reasonable not to hold the little scientific interest in this subject responsible for such matter of fact. Rather, this should be explained by the difficulties associated with the process of data collection and analysis, where the choice of representative data and finding of contradictory statements prove to be the main obstacles.

One of the first attempts to classify textual contradictions was done by Svintsov (1979). In his study on text editing, Svintsov argues that contradictions differ in the way they are expressed in text as well as in their location in the text. Based on this observation, Svintsov

proposes to distinguish between explicit and implicit contradictions on one side, and contact and distant contradiction on the other side, respectively. He further combines them into four types: contact-explicit, contact-implicit, distant-explicit, and distant-implicit.

Contact contradictions are observed if two parts of contradiction occur in the same sentence or in two different sentences directly adjoining each other. Distant contradictions, in turn, are those that have a portion of text between the parts of the contradictions or which are separated temporally.

The difference between explicit and implicit contradictions is of a structural nature. While the parts of an explicit contradiction are overtly (verbally) expressed and directly related to each other as in *Socrates is a philosopher* and *Socrates is not a philosopher* (referring to the same person *Socrates* in the real world), the parts of an implicit contradictions are not expressed directly (or verbally) and have to be inferred by the reader from other propositions in the text. For their detection, additional logical inferencing and a reader's world knowledge is required.

In some cases, however, the borderline between explicit and implicit contradictions can be provisory (cf. Svintsov 1979: 194) as shown in (3.35). Such cases Svintsov proposes to classify as explicit despite the negation that in this contradictory pair of statements is expressed implicitly (*healthy = not ill* or *ill = not healthy*). Nevertheless, the pair of statements can also be classified as an implicit contradiction.

(3.35) a. Adam is ill.

b. Adam is healthy.

Contact-explicit contradictions, which can formally be represented as *p and not p*, represent a prototypical contradiction especially when including explicit negation as in *Socrates is a philosopher and Socrates is not a philosopher* (assuming that *Socrates* denotates to the same person). The most studies, however, show (e.g. Svintsov 1979; Weimer 2005) that such kind of contradictions rarely occur in texts as these are easy to detect, and that is why they are usually corrected within the text production process. In most cases, contact-explicit contradictions arise as a result of inadvertence and negligence (Svintsov 1979: 196). Distant-explicit differs from contact-explicit contradiction only when there is a portion of text or a temporal distance between the parts of contradiction.

While contact-explicit contradictions are those that happen on the surface of the text and can be recognized without a deep meaning analysis, the recognition of contact-implicit contradictions requires additional logical inferencing and application of world knowledge. Svintsov (1979: 195) illustrates a contact-implicit contradiction by the following example (3.36):

(3.36) The first prize was given to the experienced grandmaster L. Stein who in total collected ten points (7 wins and 3 draws).

According to chess rules, a player receives one point for having won a game and a half point in the case of a draw. By performing a simple mathematical operation (7 x 1 + 3 x 0.5 = 8.5), it can be concluded that player L. Stein had only collected eight and a half points, and not the ten points stated by the author. A contact-implicit contradiction between the statements on the chess points collected appears here as the result of incorrect mathematical computation or the text producer's inadequate knowledge of chess rules.

Of all four types of contradictions distinguished by Svintsov (1979), the distant-implicit type of contradiction is the most difficult to detect. On the level of text production, distant-implicit contradictions appear to be the result of an insufficient understanding of the subject by the text producer (cf. Svintsov 1979: 197). The distant-implicit kind of contradiction can be illustrated using the example of the excerpt from Daniel Defoe's *Robinson Crusoe* that was provided in the introductory chapter of the present study.

3.4.2 Classification of Mučnik

Mučnik (1985) offers the most comprehensive classification of textual contradictions with focus on literary texts. He fully adopts the classification of logical textual contradictions developed by Svintsov (1979) and enriches it with the classification of contradictions that occur in literary texts. He refers to this kind of contradictions as abstract, or double image. Thus, Mučnik strictly distinguishes abstract contradictions from logical and regards the former to be an autonomous class of textual contradictions. As a special case of textual contradictions, Mučnik also addresses psychological contradiction. The classification of contradictions developed by Mučnik is presented in Figure 3.

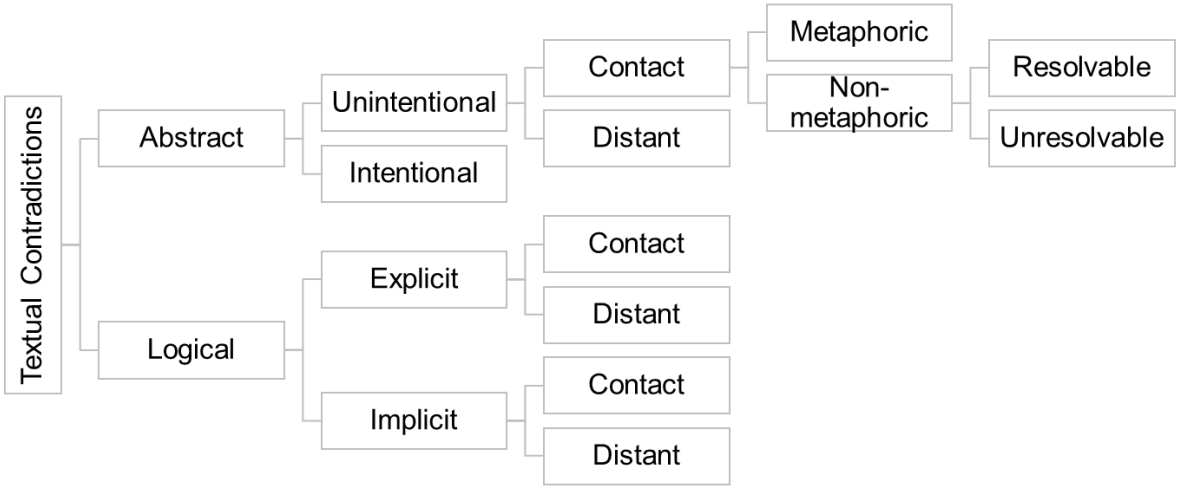

Figure 3: Classification of contradictions proposed in Mučnik (1985).

In contrast to logical contradictions, which arise in case of an incompatibility of two propositions, abstract contradictions arise as the result of an incompatibility of two mental images

as, for example, with metaphors (cf. Mučnik 1985: 148). In other words, while logical contradictions arise in cognition, that is, when the parts of an implicit contradiction can be inferred from other propositions expressed in text and are then found to relate to each other, abstract contradictions arise in the imagination and are mental representations which are incompatible with each other (Mučnik 1985: 149). The abstract contradiction can be illustrated by the following example (Tolstoy (1956) cited in Mučnik (1985: 148)):

(3.37) The glow of the sunrise played a huge flame of fire on the blue quiet sky.

In (3.37), the first part of the quotation gives a picture of a vast raging heavenly fire, and the second of a quiet blue sky. According to Mučnik following Tolstoy (1956), we observe a visual incompatibility or an imposition of an image on top of another image which, in turn, leads to a contradiction.

Mučnik further proposes classifying the abstract contradictions into intentional and non-intentional – those which are created as conscious aspiration of the author and those which arise against the will of the author, respectively. Non-intentional contradictions, in turn, are classified as contact and distant in analogy with Svintsov's classification of logical contradictions (see Section 3.4.1). Mučnik distinguishes between two kinds of contact contradictions – metaphoric (catachresis) and non-metaphoric as shown by the examples (3.38a) and (3.38b), respectively.

(3.38) a. The ship of the desert cheerfully walks along the hot sand of the Sahara.

b. He did not notice anyone nearby, except for a distant child's figure.

Finally, the non-metaphoric contradictions can be classified into resolvable – those which can be annulled by a correction of the provoked image, and unresolvable – those which keep their contradiction status because they cannot be rethought. Mučnik (1985: 160) illustrates a resolvable contradiction by an example, *A toad, lying belly up, dragged on*. The provoked image is a toad which is dragged on while lying with its belly up. This image appears to be "ridiculous, illogical" (Mučnik 1985: 161). The corrected image, in turn, would be a toad which was lying with her belly up, then flipped over before dragging on and then being dragged on. The image is corrected by the reader, and the abstract contradiction is cognitively resolved. In contrast, an abstract contradiction that arises in *House like a swallow's nest clung to the most comfortable place, it stands firmly and seems it has grown into the ground* does not provoke an image that can be rethought (Mučnik 1985: 161).

3.4.3 Typology according to Contradictory Element

3.4.3.1 In Educational Psychology

Textual contradictions were also the focus of educational psychology with its research peak during the 70s and 80s. A number of methodological paradigms have been developed during this time, which laid the groundwork for later studies.

Contradictions were, on one side, studied autonomously with a primary interest in proposing models for an explanation on how contradictions are being processed by humans and for identifying the reasons why humans fail in this task (e.g. Otero/Kintsch 1992; Kintsch 1988; Johnson-Laird et al. 2004). On the other side, contradictions were applied as an instrument in the study of text comprehension, especially for comprehension monitoring (e.g. Baker/Anderson 1982; Otero/Campanario 1990; Albrecht/O'Brien 1993; Long/Chong 2001; Wassenburg et al. 2015). The ability of contradiction, or inconsistency detection,[17] was regarded as an indicator of whether and how well the text was understood by a reader. Hereby, researchers operated with different types of contradictions.

Thus, Markman and Gorin (1981: 320) studied how children between the age of 8 and 10 years "become able to judge their comprehension, what difficulties they have, and how they can be helped to overcome these problems". The authors distinguish between contradictions arising as a violation of the background on one side, or speaking in their terms, the "commonly-held" knowledge of the reader in isolation from other statements as in (3.39a) and on the other side, as a conflict or incompatibility between statements in the text as in (3.39b) (Markman and Gorin 1981). These are referred to as falsehoods and inconsistency, respectively. Additionally, Markman and Gorin (1981) distinguish between explicit and implicit inconsistencies. In contrast to explicit inconsistencies, the recognition of implicit inconsistencies requires additional inferencing and background knowledge. However, the study of Markman (1979) showed that though implicit inconsistencies seem to be more difficult, children of 11 years have difficulties in processing both explicit and implicit inconsistencies.

(3.39) a. Corn can be served in many ways. I've never met anyone who didn't consider corn, in one form or another, one of their favorite foods. Corn can be steamed and served with melted butter; mixed with flour and egg to make bread; or made into popcorn for a favorite snack. *Most people prefer not to eat their favorite foods*.

[17] The use of the terms contradiction and inconsistency appears in the studies to be used in a loose manner. They often occur as synonyms have however to be clear distinguished from each other (Williams 1981). According to set theory and classical logic, it can be said that a set of objects (statements) is consistent, if they are all true in this set. In turn, inconsistency is observed if a contradiction can be derived from this set. That is, one statement cannot be proved to be true (=can be proved to be false) regarding other statements in the set. Also, a single object (statement), if it entails a contradiction, can be regarded as inconsistent.

b. Corn can be served in many ways. *I've never met anyone who didn't consider corn, in one form or another, one of their favorite foods*. Corn can be steamed and served with melted butter; mixed with flour and egg to make bread; or made into popcorn for a favorite snack. *The people I know don't enjoy eating corn very much*.

Garner (1981), in turn, operates with contradictions appearing as the result of lexical incompatibility between two statements as illustrated in (3.40). She uses the term informational inconsistency to refer to this kind of statements. Informational inconsistency corresponds to explicit inconsistency as defined in Markman and Gorin (1981).

(3.40) The train stopped in Centerville every day at both one o'clock and *at five o'clock*. Dr. Jones needed to travel from Centerville to Milltown on business. He decided to go by train. He packed his bags. He caught a train *at seven o'clock*, and was in Milltown in time for his meeting.

Harris et al. (1981), in their study of text comprehension among younger children, applied contradictions which appeared in the sequence of steps as the result of some steps belonging to different scripts as illustrated in (3.41).

(3.41) John is waiting. (Harris et al. 1981: 215)
There are two people before him.
After a while, it is his turn.
He sees his hair is getting shorter. (1)
Luckily, there are no cavities this time. (2)
After a while he can get up.
John puts his coat on.
He can go home.

The notion of script was developed by Schank and Abelson (1977: 41) to describe actions analogous to the Minsky's notion of the frame for objects and is defined as "a predetermined, stereotyped sequence of actions that defines a well-known situation". A script contains a variety of props, participants and their roles, the entry conditions, results, and scenes, which are referred to as headers (Hidalgo Downing 2000: 125).

Primarily, Schank and Abelson, by developing the script, aimed at elaborating on a model for the representation of human knowledge so that it could be applied in Artificial Intelligence. A starting point or idea for script was the notion of episodic memory (Tulving 1972) – a component of a long-term memory which is "organized around propositions linked together by their occurrence in the same event or time span" (Schank and Abelson 1977: 17). Taking this into account, they assumed that certain situations contain or are coded as a fixed sequence of actions, and underlying this idea, by the notion of the script. Traditionally, the script is exemplified by the situation of visiting a restaurant (3.42) provided in Schank and Abelson (1977: 39).

(3.42) John went to a restaurant. He ordered a coq au vin. He asked a waiter for the cheque and left.

In general, Schank and Abelson (1977) distinguish between three types of scripts, including situational scripts (e.g., bus, jail), personal scripts (e.g., being a friend), and instrumental scripts (e.g., starting the car). It is additionally noted that scripts are only able to deal with familiar situations and fail in case of new or unfamiliar situations.

One of the main purposes of scripts is to render it possible to recover not explicitly expressed actions in a given discourse. That is, when reading (3.42), we activate a script of the restaurant so that we are able to recognize the function of the waiter in the situation as well as to infer that John ate the food that he ordered, and also that he paid for it before he left though this information is not expressed explicitly. Moreover, scripts allow us to recognize variations in the script elements by means of so-called tracks. Examples of tracks with regard to a restaurant script can be an Italian restaurant, coffee shop, etc.

Applying this knowledge to the above example (3.41), it can be observed that the incompatibility arises between sentence (1) and sentence (2). Whereas (1) can be assigned to a script “at the hairdresser’s”, (2) is “at the dentist’s”.

In analogy with Markman and Gorin (1981), Reis and Spekman (1983) distinguish between text-based contradictions, which refer to conflicting ideas/sentences in the text, and reader-based contradictions, which arise as a conflict between text and prior knowledge as shown in the example (3.43). The example is taken from Reis and Spekman (1983: 52). Text-based contradictions correspond to explicit inconsistencies and reader-based contradictions correspond to implicit inconsistencies as defined in Markman and Gorin (1981). In turn, Baker (1985) uses the terms of internal and external to refer to reader-based and text-based inconsistencies, respectively.

(3.43) It was very warm as Don and Sue walked outside. The sun was shining brightly, and they knew the day would be a hot one. They packed the car with food, swim fins, and sun tan oil. In a short time, they would be ready to begin the ride to the beach. Before they left, they made sure they had everything. At the last moment, Sue ran into the house to get her *wool coat*.

3.4.3.2 In Computational Linguistics

The typologies proposed in the field of computational linguistics and NLP are computationally oriented. For this reason, the research is primarily focused on the cues of contradictions that can be processed by machine for contradictions to be detected.

The early studies on CD distinguished between contradictions arising from using negation and antonyms (Harabagiu et al. 2006) and developed their systems for the detection of these contradiction types only. Contradictions of the negation type, according to Harabagiu et al. (2006), include overt (directly licensed) and indirectly licensed negation. The researchers attribute *n’t* (such as in *don’t*, *can’t*, *won’t*) and *not*, negative quantifiers (*no*, *no one*, *nothing*), and negative adverbs (*never*) as belonging to the overt negation category. The

indirectly licensed negation category includes verbs or phrasal verbs (e.g., *to deny*, *to fail*, *to refuse*, *to keep from*), prepositions (e.g., *without*, *except*), weak quantifies (e.g., *few*, *any*, *some*), and NPIs (e.g., *a red cent*, *any more*)[18].

De Marneffe et al. (2008) were the first to propose a more elaborated typology of contradictions, based on the RTE corpora (see Section 2.2.2), and compiled a corpus of real-life contradictions (see Stanford Corpus of Real-life contradictions in Section 2.2.3). The researchers identified six types of contradictions based on linguistic and non-linguistic features including negation, antonym, numeric, factive, lexical, structural, and world knowledge. The contradiction types, including examples, are presented in Table 14.

Contradictions of the type negation can be recognized by identifying the negation markers, e.g., *not*, *no*, and *few*. The uses of antonyms and words with negation prefixes, such as *anti-* and *un-*, is characteristic of the antonym contradiction type. Numeric contradictions, in turn, arise as the result of numeric mismatches, including numbers, dates, and times. According to de Marneffe et al. (2008: 1041), these types are "relatively simple to detect" as in most cases, no complete sentence comprehension is required. In contrast, factive, structural, lexical, and world knowledge-based contradictions are "more difficult to detect automatically because they require precise models of sentence meaning". Factive contradictions, for example, arise from an incompatible facticity between statements or a *T-H* pair. The recognition of factivity in de Marneffe et al. (2008) is based on the recognition of factive, implicative, and non-factive verbs as summarized in the PARC lists verbs (Nairn et al. 2006). Moreover, contradictions of a factive type are recognized by identifying the modality of the *T* and *H*. De Marneffe et al. (2008) distinguish between six modalities, including possible and not_possible, actual and not_actual, and necessary and not_necessary. Thus, e.g., a *T-H* pair with modalities possible-not_possible is considered to construe a contradiction that is of a factive type. To continue, the use of oppositional terms and paraphrasing is characteristic for lexical contradictions. Structural contradictions arise as the result of an incompatibility in the syntactic sentence structure. This is observed, for instance, in a situation when the subject of the *T* overlaps with the object of the *H*. Contradictions of type world knowledge are based on discrepancies in world knowledge. Unfortunately, de Marneffe et al. (2008) do not provide any further information on this type of contradiction.

[18] Original terminology as used in Harabagiu et al. (2006).

Type of contradiction	Example
Negation	T: A closely divided Supreme Court said that juries and not judges must impose a death sentence. H: The Supreme Court decided that only judges can impose the death sentence.
Antonym	T: Capital punishment is a catalyst for more crime. H: Capital punishment is a deterrent to crime.
Numeric	T: The tragedy of the explosion in Qana that killed more than 50 civilians has presented Israel with a dilemma. H: An investigation into the strike in Qana found 28 confirmed dead thus far.
Factive	T: The bombers had not managed to enter the embassy. H: The bombers entered the embassy.
Structural	T: Jacques Santer succeeded Jacques Delors as president of the European Commission in 1995. H: Delors succeeded Santer in the presidency of the European Commission.
Lexical	T: In the election, Bush called for U.S. troops to be withdrawn from the peacekeeping mission in the Balkans. H: He cites such missions as an example of how America must "stay the course".
World Knowledge	T: Microsoft Israel, one of the first Microsoft branches outside the USA, was founded in 1989. H: Microsoft was established in 1989.

Table 14: Contradiction types distinguished in de Marneffe et al. (2008: 1041) with examples.

Later, Ritter et al. (2008) extended the world knowledge type with contradictions that arise from functional relations, such as *bornIn(x, y)*, where a person *x* is mentioned to have two birthplaces *y* though only one birthplace is possible (3.44).

(3.44) Mozart was born in *Salzburg*.

Mozart was born in *Vienna*.

Finally, Tsytsarau et al. (2010, 2011), in turn, addressed the contradictions that arise as the result of opposing opinions and feelings, which they call sentiment-based contradictions, such as statements on Internet government control as in (3.45). The example is taken from Tsytsarau et al. (2010: 1196).

(3.45) *Pro*: I suppose we better wrap a firewall around our country and not let those damn foreigners access our internet.

Contra: While it sounds like a decent idea, I'm really all for the whole uncensored and unregulated internet. I really like my internet the way it is.

3.5 Causes and Functions of Contradictions

3.5.1 Causes

There are few studies addressing the reasons for the contradictions arising in text and speech. Traditionally, the text of *De Sophisticis Elenchis* (Sophistical Refutations) of Aristotle's *Organon* is considered to be the main work on this subject, followed by the studies of the Middle Age philosophers Thomas Aquinas and William Sherwood. A detailed overview of these studies is provided in Rieger (2005). The causes of contradictions have also been addressed in Mučnik (1985), Weimer (2005), and Govier (2010).

Aristotle, as stated in Rieger (2005: 95ff.), holds fallacy as a possible reason that leads to a contradiction. A fallacy is defined as an error that arises as the result of incorrect logical reasoning by constructing a valid argument. According to other (more focused) definitions, it is a conclusion which does not follow from the explicitly expressed or implicitly presumed premises as a result of an incorrect application of the logical rules of reasoning (Apel 1976; Blackburn 2016). It should, however, not be assumed that the conclusion is necessarily false. A fallacy does not declare the truth value of the conclusion. Moreover, fallacy can be also observed in case of a valid conclusion which, however, follows from wrong premises.

Aristotle lists in his *Sophistical Refutations* 165b and 166b a total of thirteen fallacies that can cause a contradiction (Rieger 2005). He divides them into fallacies in the language (in dictione) and fallacies not in the language (extra dictionem). There are, in total, six fallacies in dictione which include equivocation, amphibology, combination, division, accent, and form of expression, and seven extra dictionem fallacies, including accident, secundum quid, ignoration elenchi, petition principia, non causa pro causa, consequent, and many questions (Parry/Hacker 1991; Rieger 2005). The definitions of these fallacies are summarized in Table 15.

Fallacy	Definition / Example
Fallacies in the language (caused by ambiguities in the language)	
Equivocation	*Definition:* Arises in case of using an ambiguous term or a phrase. *Example: All bark is grown on trees. – All dogs bark. – All dogs are trees.* *Remarks:* Modern logicians provide the same definition of equivocation.
Amphiboly	*Definition:* Arises as the result of a faulty grammatical construction of a sentence so that the latter can be interpreted in multiple ways. *Example: One morning I shot an elephant in my pyjamas.* *Remarks:* The same view on amphiboly is shared by modern logicians.
Combination (composition by Aristotle)	*Definition:* Arises as an improper combination of words. Combination as defined by Aristotle differs from view of modern logicians on it. According to the latter combination is observed "when the same term is used divisively in a premise and collectively in the conclusion". *Example: All dogs are common. / All albino spaniels are dogs. / All albino spaniels are common.*

Division	*Definition:* According to Aristotle an improper division of words. *Example: 5 is 2 and 3, therefore 5 is 2 and 5 is 3* (2 and 3 correctly combined in the premise are incorrectly divided in the conclusion). *Remarks:* According to modern logicians, fallacy of division is observed either when "the same term is used collectively in the premises and used divisively in the conclusion" or "a term is predicated of a class taken collectively in the premises, and in the conclusion is predicated of something said to be a member of this class". *Example: The French are a great nation. / Charles de Gualle is French. / Charles de Gaulle is a great nation.*
Accent	*Definition:* Ambiguity which arises as the result of changing the stress in the word meaning a contrary to what a sentence says. Such cases are frequent in Greek, but rare in English. For this reason, accent for modern logicians means also as sentence stress or emphasis on a word or phrase.
Form of expression (misleadingly called figure of speech)	*Definition:* Fallacy of form of expression refers to the case when it is argued that words and sentences that are grammatically similar are also logically and semantically similar. *Example: "[…] "flourishing" is a word which in the form of its expression is like "cutting" or "building" [as in building a house] yet the one denotes a certain quality – i.e., a certain condition – while the other denotes a certain action" (Sophistical Refutations 166b15).* *Remarks:* Ignored by the most logicians.
Fallacies not in the language (no common cause)	
Accident	*Remarks:* s. secundum quid (a dicto simpliciter ad dictum secundum quid).
Secundum quid	*Definition:* Includes a) the fallacy of introducing qualification, or a dicto simpliciter ad dictum secundum quid and b) the fallacy of eliminating qualification, or a dicto secundum quid ad dictum simpliciter (b). *Definition (a)*: A fallacy which arises from unqualified use of a term as if it is qualified. *Example (a): Of course I'm promiscuous, Jesus said we should love our neighbour.* *Definition (b):* A fallacy which arises from qualified use of a term as if it is unqualified. *Example (b): Steel in the form of penpoints is not a good heavy construction material, hence steel is not a good heavy construction material.*
Ignoratio elenchi (ignorance of the refutation, irrelevant conclusion, or conclusion-thesis gap)	*Definition:* Valid or invalid conclusion which however does not refer to the question in issue. "Arguing beside the point". *Remarks:* The same treatment of the fallacy by modern logicians.
Petitio principia (begging the question)	*Definition:* Assuming in the premises what needs to be proved. Can be constructed by different means including argument in a circle, question-begging terms, question-begging questions. *Example: He ought to get a raise, because he should have a higher salary.* *Remarks:* Same definition by modern logicians.

Non causa pro causa (false cause)	*Definition:* Aristotle understood it as a contextual fallacy which is seldom for contemporary argumentation. Modern logicians understand it as a cause which is mistakenly held as a cause for some event. *Example: After I wore Aunt Maud's tie I received an "A" on the physics exam, therefore it's a lucky tie* (false cause: wearing a tie of Aunt Maud on the physics exam).
Consequent (also illicit converse, affirming the consequent)	*Definition:* False (or invalid) conversion. *Example: If B is the consequent of A, then non-B is the consequent of non-A.* *Remarks:* Seldom mentioned by modern logicians.
Many questions	*Definition:* Arises when two or more questions are asked at once. *Example: The execution of Mary Queen of Scots was brutal and sacrilegious – was it, or was it not?* *Remarks:* In modern logic is regarded as the result of assumptive questions such as *Why must all democracies be corrupt?*

Table 15: Definitions of Aristotle's fallacies as provided in Parry/Hacker (1991: 423-457).

Besides fallacy, insolubilia, or insolubles, is another possible cause of contradictions as stated in the De insolubilibus logical tractate by William of Sherwood (Rieger 2005). In the Middle Ages, the term insolubilia referred to the term more known at present as the liar paradox and its variations. For more information on insolubles, see, i.a., Rieger (2005: 115ff.). For the relation between the concepts of contradiction and paradox, see the previous Section 3.1.4.

Mučnik (1985), in turn, holds different kinds of stylistic error responsible for contradictions arising in text, including the incorrect understanding/assignment/interpretation of syntactic roles – subject and object (3.46a), compositional ambiguity (3.46b), displaced sentence emphasis (3.46c), the use of an odd word/phrase/clause, repetition (3.46d) as well as double imaging (3.46e).

(3.46) a. Leto smenjaet osen'. (Mučnik 1985: 115)
Summer-replaces-autumn.

Similar to English, Russian is an SVO language but with more flexible word order. Thus, commonly *leto* (summer) as it is located in first place would be interpreted by the reader as subject and ***osen'*** (autumn) as object. This, however, is cognitively perceived as a contradiction by the reader as autumn comes after summer and not vice versa. By repeated reading, the assignment of syntactic roles is rethought and the assignment of syntactic roles is corrected so that the ***osen'*** occurs as a subject and *leto* has the role of a direct object.

b. She met the man with her friend. (Löbner 2013: 48)

Is the PP *with her friend* related to the verb *meet* meaning *she and her friend met a friend* or is an attribute of the NP meaning *the man who was with her friend*?

c. Some writers have a portrait in more detail, others – a speech characteristic. (Mučnik 1985: 91; translation from Russian – N.K.-B.)

> During the first reading, the emphasis is perceived by the reader in *in more detail*, while the author thought it for *portrait*. After the second reading, the reader corrects this perception error.
>
> d. From the 5th to the 15th of May, the hairdresser *will be closed* for repairs, the hairdresser *will not be working*. (ibid., 36; translation from Russian – N.K.-B.)
>
> e. The glow of the sunrise played *a huge flame of fire* on the *blue quiet sky*. (ibid., 38; translation from Russian – N.K.-B.)

The causes for the contradictions appearing in human conversation, in turn, are discussed in Weimer (2005) and also partly addressed in Govier (2010). Weimer (2005) provides a list of four causes, including real-life examples, pointing out that although the list is in no case complete, it tends to represent the most frequent cases. The causes identified by Weimer (2005) are the following:

- A speaker is not aware of rules, requirements, and assumptions that underlie his statements;
- A speaker is not aware of how his statements, which are generally formulated about other statements, also refer to themselves;
- A speaker is not aware that his aims and the actions he takes are incompatible;
- A speaker does not remember what he said previously, or he does not remember what he previously said to another person. This cause of contradiction in a conversation is also addressed in Govier (2010: 52).

3.5.2 Functions

According to the traditional view, contradictions which the result of a violation of the laws of logic signal errors in cognition/reasoning and have, therefore, to be avoided. In logic (with the exception of representatives of paraconsistent logic), this aspect is treated under the term triviality, claiming that everything can be derived from two contradictory sentences. This principle is also known as ex falso quodlibet, or the ex falso sequitur quodlibet notion.

In semantic theory, the view on contradictions is of a negative nature as well. As regards two statements, of which one is false and the other is true without further knowledge about which one is true and which one is false, as well as considering that the overall truth value of the conjunction of these two statements is false, Levinson (1983: 194) and Weimer (2005: 99) conclude that contradictions are meaningless and uninformative since they do not satisfy the human need for information.

But are all contradictions occurring in natural languages “bad” and need to be avoided?

One of the most representative kinds of an intentional “good” use of contradiction is the production of a humorous effect, irony, and absurdity. A prominent example of the usage of contradictions for the creation of absurdity is represented by Lewis Carrol’s *The Adventures*

of Alice in Wonderland. Humor and the role of contradiction for humor creation both were the subjects of many studies such as Freud (2009), Norrick (1986), Schulz (1976), to mention only a few.

The humorous effect based on contradiction can usually be achieved by several techniques. These include first the creation of a cognitive conflict between an expectation of the recipient/reader and final realization of this expectation, or real outcome. Consider the following example (3.47), which is usually attributed to Twain, though the original joke belongs to Ebenezer R. Hoar:

(3.47) I did not attend his funeral, but I sent a nice letter saying I approved of it.

Second, the humorous effect can be achieved by means of lexical, structural, and phonetic (in spoken texts) ambiguity – contrast between what is said and what is meant (3.48). The realization of these techniques is typically applied in anecdotes/jokes (Raskin 1985):

(3.48) How do you make a turtle fast?
Take away his food.

In addition to the creation of a humor, contradictions are often used intentionally as a means for achieving an emphasized expressiveness and special effect in literary texts and especially in poetry as exemplified by a short poem of the Roman lyric poet Catullus (3.49):

(3.49) I hate and I love
Why do I, you ask?
I don't know, but it's happening
and it hurts.

The contradiction is realized here as an oxymoron, which is an intentional connection between two opposite concepts. Other figures of speech which are based on contradiction are catachresis (the use of words and phrases in a not-traditional, unconventional way) and antithesis (“a seeming contradiction of ideas, words, clauses, or sentences within a balanced grammatical structure” as defined in the Columbia Encyclopedia (2017).

Beside literary texts, contradiction often occurs in proverbs and quotes. The most famous quote is Socrates’ *I know that I know nothing*, which is also known as the Socratic paradox.

Furthermore, the contradiction can be used as a means to gain the reader’s attention in order to convince him to begin or continue to read the text. Some psychological studies (e.g., Campion et al. 2009) also showed that contradictions occurring in the text as a source of uncertainty raise the cognitive interest of readers, making them read further. For this reason, to achieve this effect, contradictions are often already located in the title of the story or a movie or in news headlines, and are usually realized as an oxymoron, or paradox (Ganeev 2004), such as the movie title *She’s a man*.

To conclude the discussion of the intentional use of contradictions in speech and text, it should be mentioned that contradictions are often used in debates as a rhetorical means for the purpose of irritating and manipulating the conversational partner (an intentional “bad” use). Thus, the ability to identify logical fallacies in the arguments of others is valuable in order to save oneself from manipulation by rhetorically skilled speakers.

3.6 Summary

In the present chapter, contradiction as one of the key concepts of the present study was introduced into the framework of logical theory as well as with regard to its use in language. First, three views on contradiction – the ontological, the doxastic, and the semantic – have been presented as elaborated by Aristotle (traditional logic). The semantic view, which defines contradictions by means of truth values, will not be further considered for conceptualization and implementation in a CD system as at present, the machines are not able to determine the truth value of the propositions. In contrast, the ontological view according to which a contradiction is observed in case the two propositions satisfy five conditions, namely, (a) reference to the same thing, (b) expression of the same proposition about this thing, (c) the same time reference, (d) presence of negation as sentence operator, and (e) exclusive and exhaustive disjunction, will be prior to a system implementation. Besides the ontological definition, also the semantic view on contradiction, which states that it is impossible to believe a thing to be and not to be, will be considered for the development of a CD system as described in Chapter 7.

Second, besides contradiction, this chapter presented the concepts of contrariety, paradox and tautology, which are related to contradiction. Due to their similarity, contradiction and contrariety are often treated as synonyms. Though also in the present study, these concepts will be addressed jointly as contradiction, it should be considered that while both are exclusive, contrariety, in contrast to contradiction, is not exhaustive. Regarding the concepts of contradiction and tautology, the difference between them is that the latter is always true while the former is always false in every possible interpretation. In turn, the difference between contradiction and paradox is that the latter is both true and false at the same time.

Third, besides the definitions of contradiction and their related concepts, it was shown that natural language negation can be expressed by a variety of means and is not limited to the negation operator not. This elaboration on negation is focused on the English language.

Further, the status of contradiction has been discussed under the phenomena of modality and vagueness, by considering the multiple word meanings and the context of utterance as well as by fixing the references. In the context of vague language and context of utterance (CoU), the concepts of a borderline contradiction and fake contradiction, respectively, have

been introduced. It was shown that the presence of contradiction is, in many cases, a matter of interpretation, which by this means poses a challenging issue for a CD task. In case of modality, there are still discussions about what is the best approach for handling contradictions.

Finally, the chapter addressed the possible causes and functions of contradictions by claiming that not all contradictions have to be avoided, indicating the role of the reader in the final decision on whether a contradiction has to be eliminated or not.

4 The Characteristics of News Texts

Moving away from contradiction, Chapter 4 introduces the second key concept of the study, which is news articles with a focus on online news. We believe that the knowledge about the particularities of online news texts can contribute to a better implementation of the CD task. The chapter begins with basic notions related to news articles, including the definition of news as text genre, a discussion of differences between online and print news articles as well as the values which drive the process of news production (Section 4.1). The elements and structure of (online) news article are the topics of Section 4.2. The characteristics of news language are presented in Section 4.3 and include i.a. reported speech, event categories, numbers, and figures as well as time and place mentions. Though the information in online news articles can be communicated in multiple ways besides text, in most cases including also images, video, etc., only the textual mode will be taken into consideration for the purpose of the study.

4.1 Introductory Notions

4.1.1 News as Text Genre

There exist a number of definitions of *news*, which are inconsistent and superficial (Lage 1979, cited in Bonini 2009), leading to a terminological confusion. Bonini (2009), following van Dijk (1988), relates this problem regarding the ambiguity of the term news. According to van Dijk (1988: 4), the notion of news may be understood (a) as new information about some events, persons, or things, (b) as a news item or report, i.e. a text or a discourse on TV, radio, or in a newspaper, (c) and also a as TV or radio program in which news or new information is presented. In order to avoid terminological confusion in the present study, the term news will be used in the sense of “news as fact, some new information”. By referring to the news in the sense of a news item or report, we will use the term news article (text along with images, video, etc.) or news text (plain text only). In case of usage in another sense, the term will be specified accordingly.

News (both in the sense of new information and news articles) can be provided through different media such as print, online, broadcast, and TV. Online media include, in turn, besides the websites of newspapers and news agencies, also other media channels such as e-mail, mobile phones, and social networks. Traditionally, however, news articles are treated along with printed newspaper and at present, also in the context of an online newspaper.

In general, the texts of a newspaper, according to their communicative function, can be classified into informational, opinion/persuasive, instructive, bicentric-reciprocative, and

contact-building texts[19] (Lüger 1995). News articles, together with, e.g., reportages and weather reports, belong to the informational texts with the primary purpose of informing the reader. Based on their length, news articles can further be classified into a news item which is a shorter news article and news report which is a longer one.

Depending on the topic of the events presented in the news article, one distinguishes between soft and hard news articles. The hard news refers to information on topics of "high importance" such as national and international politics, business, health, and education. Soft news, in turn, relates to topics of human interest such as entertainment, celebrities' lives, including scandals, lifestyle, sports, arts, and culture (Lüger 1995: 103). Also, news on natural disasters and crime are ascribed to soft news.

Besides the thematic difference, soft and hard news exhibit differences in purpose, language, and structure. In contrast to hard news, the main purpose of which is to inform the reader objectively, the soft news pursues a goal to win over, or canvass a reader (Lüger 1995, 108 used the German term lesewerbend to describe this function of soft news). That is, the language of soft news is characterized by the usage of vivid expressions and figures of speech, contrasting with the objective and neutral language of hard news. Regarding the structure, soft news exhibits an emphasis on the introductory and closing parts and the chronological order of the events, while hard news follows a fixed structural schema, placing information from the most important to the less important (Lüger 1995). The structure and elements of news texts are described in more detail in Section 4.2. The description provided can, however, be mainly applied to the hard news.

The production of a news article is processed in a "complex and often cyclical route", involving "many hands" doing the work (Bell 1991: 34). According to Bell (ibid.), the procedure of news article production, based on the example of a Medialink copy system, includes the following stages: Starting at the chief reporter's desk, with the information/or a certain document such as, e.g., a report, the story is then assigned to a journalist for elaborating the story, that is to search for related background information, earlier stories, conducting interviews etc. When the story is ready, it again lands on the desk of a chief reporter whose task it is to check the story for poor writing, content gaps, etc. The ensuing corrections are then integrated directly by the chief reporter or sent back to the journalist. After the first changes have been made, a subeditor is the next to edit the article including "cutting, tightening,

[19] The terms have been freely translated from English into German by the author of the present monograph. The original German terms are informationsbetonte Texte, meinungsbetont-persuasive Texte, instruierend-anweisende Texte, bizentrierte Texte and kontaktherstellende, respectively. Besides Lüger (1995), the classification of different text genres of newspaper has been also addressed in Weischenberg (2001), La Roche (2006), Reumann (2002), and Schneider and Rauer (2009), only to name a few.

clarifying, reordering, and restyling" and checking the copy for inconsistencies (Bell 1991: 35). In case of gaps identified, the copy is again sent to a journalist or chief reporter. The final check of the copy is performed by the editor. In the last stage of news production, the copy is transmitted, usually by the sub-editor, into a copy system such as Medialink.

It should be noted, that this procedure may have some slight differences for news articles produced in domestic journalistic outlets and for the production of online news. Thus, the latter often contain, for instance, the corrections made by the readers. Moreover, in the production process of online news articles, some of the previously described steps are often simply skipped or left out so that the news articles may be published immediately (Fenton 2012: 561).

Journalists make use of different kinds of input material in the process of news text production. According to Bell (1991), most of the stories originate from interviews. Van Dijk (1988: 128), in turn, based on an analysis of a sample from Dutch newspapers, identified twelve kinds of input materials, including among others, dispatches of national and international news agencies, press releases, press conferences, documents of different types produced by legislative bodies, committees, and organizations (e.g. reports, letters, agendas), interviews (by telephone and face-to-face) and notes of interviews, phone calls, and press conferences.

It is interesting to note that journalistic outlets prefer the news articles that originated in other organizations, which they will then simply modify (Bell, 1991). News agencies are considered to be the prevailing choice for news articles. As the first instance in the news production and its dissemination to a variety of clients, the news material provided by news agencies is characterized by a less controversial language. However, as MacGregor (2013) and Johnston and Forde (2011) point out, the news agencies and their reporters are usually not attributed in the modified articles so that they – as the original source of the news – remain unknown to the news recipient.

Agence France Presse (AFP) is considered to be the oldest news agency, which in 1835, was initiated as Agence Havas by a Parisian translator and advertising agent, Charles-Louis Havas, and still exists today (Broderick/Miller 2007). The news agencies which dominated in the 20th century include "The Big Four" – Reuters (International), Associated Press (USA), Agence France Press (France), and United Press International (USA) – as well as TASS (the Soviet Union, today the Russian Federation) and Xinhua (China). The Big Four provided, e.g., over 90% of foreign news printed by the world's newspapers. The domestic news mainly originates from national news agencies such as, e.g., the Press Association in Great Britain and New Zealand, and the Associated Press in the USA (Bell 1991: 16-17).

4.1.2 Online Newspaper vs. Printed Newspaper

Online newspapers and news articles have been studied from their structural, linguistic, and psychological perspectives, e.g., in Ihlström and Lundberg (2004), Maier (2005), Stuart (2006), Opgenhaffen and d'Haenens (2011), Hoffmann (2016), Ohler (2016), and Ohler and Schwiesau (2016), only to name a few.

The prevalent view on online newspapers (also on digital newspapers and Internet newspapers) is that they represent the "web-editions of print papers" (van der Wurff 2005: 4). A number of studies (e.g., Bednarek/Caple 2012) point out that print and online newspapers share a number of commonalities such as a standard news language and a structural similarity of news articles. However, as they are offered through different media, print newspapers (which are paper-based) and online newspapers (which are display-based) cannot be equated offhand and thus have to be distinguished (Bednarek/Caple 2012). The main distinguishing characteristics of online newspapers, in comparison to print newspapers are considered by researchers to be particularly the Internet-specific features, such as interactivity, the use of multimedia, and hyperlinking (Schwiesau/Ohler 2013; van der Wurff 2005, 2016; Fenton 2012).

Interactivity, according to van der Wurff (2005), is the most studied characteristic of online newspapers. Accordingly, Massey and Levy (1999) distinguish between the content and the interpersonal interactivity.

Content interactivity refers to the degree to which the readers are provided with opportunities to independently select the news content to read (Massey/Levy 1999; van der Wurff 2005). Content interactivity can be achieved in online newspapers by providing the news article with internal links (links to information provided by the same source), external links (links to information provided by other sources) as well as more complex elements such as search options and news archives (Ha/James 1998; Jankowski/van Selm 2000; van der Wurff 2005). Moreover, the reader is not only provided with the possibility of selecting the news content that he prefers, but also the order of the news content to appear on the computer screen by, e.g., clicking on the hyperlinks. In contrast, the order of news articles in printed newspapers is a decision made by a chief editor and is therefore fixed. Considering these aspects, van der Wurff (2005) argues that the idea of content interactivity is similar to that of databases contributing to the non-linear reading of news (Massey/Levy 1999; van der Wurff 2005).

Interpersonal interactivity, in turn, "expresses the logic of conversation" (van der Wurff 2005: 18) between a reporter and a reader(s) or/and between the readers (van der Wurff 2005). Interpersonal interactivity can be realized in online newspapers by diverse means, including providing the contact information of the reporter (e.g., an e-mail address), the possibility of

commenting on the article, or discussing it in chat rooms or discussion forums (Jankowski/van Selm 2000; Kenney et al. 2000; van der Wurff 2005).

The use of multimedia – another distinct characteristic of online newspapers, in comparison to the printed media – contributes to interactivity and precisely to content interactivity as well. The reader is provided with the possibility to choose between different formats of information reporting, which are not limited to text and graphics but can also include audio and video clips. The extent to which the online newspapers make use of different multimedia formats, and in particular audio and video, depends on technical factors such as the capacity of the server operated by the online newspaper as well as the Internet speed.

In addition to the interactivity and the use of multimedia, online newspapers also differ from the print newspapers in terms of news immediacy warranty. Enabled by the Internet speed, online newspapers can report on events and provide their updates in real-time. The latter aspect is especially advantageous as this allows the newspapers to inform their readers on incorrect facts occurring immediately after their detection, in contrast to printed newspapers which have to wait until the release of the next issue. This is one of the reasons why online newspapers are more preferred by the readers. Unfortunately, the immediacy of reporting can also have negative consequences. As already mentioned previously, immediate release of the news articles is often possible at cost of the news quality as reporters often do not have enough time for editing the text (O'Sullivan 2005; Fenton 2012).

In addition to the Internet-specific features, online newspapers differ from the print newspapers in respect to the organization of information. Besides hyperlinking, Schwiesau/Ohler (2013) mention a teaser to be a typical element of an online newspaper. The teaser is located on the start page of the newspaper and serves the purpose of providing the reader with an overview of the reported events and attract him for further reading. A teaser can be formulated as a question or an announcement or provide a summary of a reported event (Schwiesau/Ohler 2013).

Additionally, Bednarek and Caple (2012) show that print and online newspapers differ with respect to the organization of news stories. Thus, news articles on different topics in online newspapers are separated by tabs whereas the separation of articles in print newspapers is spatial with labels assigned to them.

Finally, some researchers observed (e.g. Knox 2010) that articles in online newspapers are shorter than in the print newspapers. Thurman (2007) explains this in that the purpose of the online news articles is to inform the reader in a shorter time. Another explanation of this observation can be the fact that many online newspapers by reason of being free of charge and therefore less profitable than their corresponding print version publish short articles to canvass the reader and make her/him to purchasing an online subscription or a print

subscription for more detail on the reported event. The significance of shorter online news texts for the reader and how short news texts can be critically read have been studied in the framework of systemic functional linguistics (e.g. Bateman 2008; Knox 2010).

4.1.3 Values of Selection and Production of News Articles

Before the actual process of news article production, either print or online, a decision on what information or event is worthy to be reported has to be made by a reporter and editor. This process is not arbitrary but is rather governed by a set of particular criteria, rules, and principles. In the news practice, these are referred to as values.

However, the term value in the journalistic literature is not always used homogenously (Bednarek/Caple 2012). Depending on which aspect is focused on during the news text production and besides denoting the criteria and rules for deciding which information can be news (Bell 1991; Richardson 2007; Brighton/Foy 2007) and judging which news event features are more newsworthy to be reported (Galtung/Ruge 1965/2016; Hartley 1982; Fowler 1991; Tunstall 1996), the value also occurs in two more meanings as preference and as quality. Whereas, value in the meaning of quality considers the perspective of either the news producer and news recipient (Galtung/Ruge 1965/2016; Hartley 1982; Tunstall 1996). It refers to the newsworthiness of the information from the perspective of the expected news recipient if used as a preference (Richardson 2007). Further, we refer to value as quality when speaking about that which is necessary to make a story newsworthy (Cotter 2010).

Since the study of Galtung and Ruge (1965/2016), a number of catalogs of news values have been proposed (e.g., the foundational studies of news values by Galtung/Ruge 1965/2016; Bell 1991; Fowler 1991; Richardson 2007; Cotter 2010). Despite assuming different definitions of the term value in constructing the catalogs, the latter nevertheless exhibit a number of commonalities. It should also be noted that developed for the printed and broadcast news, the values are doubtlessly also valid for online news. The most comprehensive summary and comparison of the existing catalogs are provided in Bell (1991) and Bednarek and Caple (2012).

Bell (1991) distinguishes between three kinds of values for developing a news text: values in the news text (the necessary characteristics of news), values in the news process (rules for the news text production), and values in the news actors and events (criteria for the selection of events and actors involved in these events for reporting in the news). Bednarek and Caple (2012) however argue that only the third kind of values – values in the news actors and events – are true news values, while the other two refer to the writing objectives and news market factors, respectively, and can barely be described as values.

The values in the news text, according to Bell's classification (1991: 160), include brevity, clarity, and color. As already mentioned, Bednarek and Caple (2012) define these aspects

as news-writing objectives rather than the values or general characteristics that a story has to possess in order to be taken into consideration. In addition to brevity, clarity, and color, other studies also ascribe objectivity, balance, precision, and accuracy to values in the news text (Richardson 2007; Cotter 2010).

The aspect of objectivity had developed as a central value of professional journalism at the end of the 19th century and shares much with the value of objectivity in scientific writing (Tuchman 1980: 203). That is, the ideal realization of the objectivity value is observed when the reporter distances himself from his personal, subjective point of view and creates/delivers "accurate textual copies" (value of accuracy) of an event (Schiller 1981: 87, cited in Cramer 2011: 63). Moreover, the reporter has to present different points of view on the event while ensuring that they appear to be in balance and consistent with each other and do not include mistakes, errors, or misinterpretations (value of precision).

Bell (1991: 159-160) lists continuity, competition, co-option, composition, predictability, as well prefabrication, as values in the news process, the aspects which impact the selection of events to be reported. Due to the irrelevance for the purpose of this study, these values will not further be presented in any detail.

Finally, the values in the news actors and events, as defined in Bell (1991: 156-158), include negativity, recency (timeliness in Bednarek/Caple 2012), proximity, consonance, unambiguity, unexpectedness, superlativeness (threshold in Galtung/Ruge 1965/2016), relevance (van Dijk 1988; impact in Bednarek/Caple 2012), personalization, and eliteness (also prominence in Bednarek/Caple 2012). Further values for the selection of the news actors and events are defined, e.g., in Tuchman (1980) and van Dijk (1988).

The value of negativity is related to the negative side of events reported in news and is considered as a "the basic news value" (Bell 1991: 156). Due to negativity, a negative event is likely to be reported. Indeed, it is obvious that negative events, such as accidents, wars, conflicts, disasters, death, damage, and injuries, prevail in the news. Negative news is considered to be more newsworthy, and is explained by the strong interest of the readership. Nevertheless, positive news also occurs in the news articles. Feez et al. (2008), in turn, argues that newsworthiness can be achieved by reporting both "stabilizing" and "destabilizing" events.

The value of *recency* is related to the time period required for a news story to be reported. The idea underlying this value is that the news that has just happened or had happened within the last 24 hours is valued higher and is more likely to be reported. Thus, for example, the information on a murder is valued higher than a report on an ongoing police investigation. Analogously, the verdict is valued higher than the trial.

According to the value of proximity, the selection of news is first determined by the geographical distance of the event and second, by cultural nearness (or meaningfulness, as defined in Galtung/Ruge 1965/2016), i.a., the familiarity and similarity of one country to another. Thus, an event in a small town can be newsworthy only for this town, until it has attained a more global impact.

Another value for the selection of events and actors is consonance. Following the consonance value, a news event is newsworthy when it fits the stereotypes or expectations of the audience about a particular event or actors (individuals, organizations, institutions, countries) involved in this event. In order to explain this, Schank and Abelson (1977) developed the concept of a script. According to their theory, all events have a "typical pattern" that they follow, and humans have a mental script of how particular events should happen/develop.

According to the value of unexpectedness, an event is newsworthy when it is unpredictable and rare, in contrast to the routine which does not carry any surprising charge. Thus, new kinds of information are valued higher than those which are already known. It should be noted that scientific information usually has a low value unless there are some unexpected, surprising findings to report on.

Also, stories with intensified or maximized content (a large number of participants involved, the extended scope of consequences of the event, its intensity, its size, etc.) are the preferred (the value of superlativeness) news to report on. "Generally speaking, the news value of superlativeness says that the bigger, the faster, the more destructive, the more violent, the more famous…the more newsworthy something is." (Bednarek/Caple 2012: 44)

The selection of the information source is governed by the value of eliteness. According to this value, the newsworthiness of a story is determined on one side, by the prominence of the individuals, organizations, institutions, or nations involved in the event. That is, the news about celebrities is considered to be more newsworthy than that only sourced on ordinary people. On the other hand, eliteness is determined by the prominence of the information source. Highly ranked institutions or authorities are more likely to be cited than a lesser ranked source. "The more elite the source, the more newsworthy the story" (Bell 1991: 192). The elite source is an ideal source of news because it is said that they have a special expertise or unique knowledge about issues of public concern. Also, their ability to impact large groups of people is considered to play a role (Roshco 1975: 74-75). That is, politicians are considered to be the most preferred source of information. It should be noted that the source of information can either be attributed or unattributed in a news article (Section 4.3.1). The study by Bell (1991) shows that sources with an affiliation to some institution or organization are more preferred for citation.

Further values include relevance (extent of the event's influence on the lives of the audience), which was first mentioned in van Dijk (1988), unambiguity (the use of clear facts and minimal use of "if's, but's and maybe's" (Bell 1991: 157)), facticity (frequent use of numeric and factual information, location, and person names) and personalization (personalized stories, e.g. witness accounts, attract more interest and increase the value of news more than general concepts).

As it can be concluded, the selection of the events and actors follows the principle that "the dominant, the conventional, the concrete, the rare, the extreme and the negative" all have a high value. Accordingly, news about prominent persons behaving unconventionally or performing an excessive and scandalous act is very much likely to be reported. In this respect, Cramer (2011: 69) fairly notes that values that guide the selection of news actors and events conflict with the value of objectivity. "While the objectivity promotes an informational register in the effort to eliminate standpoint, news values [of selection of actors and events to be reported – N.K.-B.] describe a professional system of preference for the reporting of certain kinds of events and the highlighting of certain features of those events." (Cramer 2011: 70).

Moreover, much research has been done (e.g. Bednarek/Caple 2012) to establish the impact of news values (the third kind of news values according to Bell's classification) on the language of news articles and precisely on the choice of particular linguistic constructions and lexical units. As the language of news is an essential basis for the CD system, it will be discussed in more detail in Section 4.3.

4.2 Structure and Elements of News Articles

The structure of news articles from the linguistic and journalistic perspective was described in, e.g., van Dijk (1988), Bell (1991), Ungerer (2004), Feez et al. (2008), Cotter (2010), and Lamble (2011). These studies focused predominantly on the text of hard news. The relation between image and text in the news article has been thoroughly studied in, e.g., Bednarek and Caple (2012). According to the research, the news text consists of four elements, including a headline, an attribution, an intro/lead, and the actual story – the body/lead development (Bednarek/Caple 2012). With some exceptions, for instance, depending on the medium of the news production and reporting (online, print), and the individual decision of the journalist, the order and realization of these elements can vary in different news articles.

The news story begins with a headline which serves several purposes (van Dijk 1988: 36; Bell 1991; Conboy 2007: 13; Bednarek/Caple 2012: 100). It attracts the reader's attention, provides a reader with a brief summary of a news story, and provides a stance towards it. Moreover, headline serves the purpose to maximize the newsworthiness of the story

(Bednarek/Caple 2012: 100). Conboy (2007: 13) additionally argues that the headline can contribute to an increasing marketization of a newspaper. After reading a headline, a reader has some expectation of the content which follows and tends to connect the content with the headline (Fries 1987: 61). One should note that headlines do not include an evaluation nor any background information. Further, the structure of headlines varies across languages (Kniffka 1980). Moreover, the content and style of the headline are the first indicators of the values of a newspaper and the way the newspaper communicates with its audience (Conboy 2007: 13).

The linguistic means related to the functions of headlines, as summarized in Bednarek and Caple (2012: 101), involve strong, "intense", emotional words; word and sound play (Lennon 2004; Caple 2010); metaphor, idioms, alliteration, proverbs, pseudo-direct quotes; omission of determiners; omission of verbs or usage of verbs without auxiliaries; the usage of verbs in the present instead of in past tense; rare use of attribution; rare specification of time – how and where rather than when; and the use of premodified noun phrases. It should be noted that although these features are typical for a news article, some of them may not occur, however. For more elaboration on the linguistic characteristics of headlines, see Mårdh (1980), Bell (1991), Reah (2003), Feez et al. (2008), Ifantidou (2009).

The study of Schwiesau and Ohler (2013) reveals that in contrast to the headlines of print news articles which, according to the study conducted by Mårdh (1980: 88), consist of an average seven words, the headlines of online news articles are shorter, consisting of only two to three words. It is interesting to note that headlines are often constructed by other news workers and not by the author of the news story himself (Bell 1991). A preferred method for finding the best headline for an online article is to create a number of headlines and then publish them online. The headlines that collect the most clicks are considered to be the best variant. Bednarek and Caple (2012) additionally point out that the headlines of online news articles are often adapted to the demands of search engines. The researchers, however, do not provide information on the exact principles from journalistic practice that have to be followed in the creation of headlines for online newspapers.

For the headline of print and online news articles, is common that the former is followed by an attribution and a lead. The attribution provides information on the source of the news (news agency, journalist byline). Additionally, it contains the news settings such as the time and place. In some cases, an attribution or its parts can be located at the end of the news article.

Lead, in turn, is a summary of the news story, without repetitions, and is a part of the news story. It is characterized by a high information content, expressing the main point of a story, its major topic, in a pair of sentences (van Dijk 1988: 53) by "raising a colourful detail"

(Cappon 1982: 31 cited in Bell 1991: 183). With a slight exaggeration, Bell (1991: 176) formulates the idea behind the lead as follows: “it compresses the values and expertise of journalism in one sentence”. The lead includes the journalistic *who*, *what*, and *where* and can be itself regarded as a micro-story (Bell 1991). The main purpose of a lead, besides a summarizing the news story, according to Conboy (2007: 11), is to increase the marketization of a newspaper.

The main requirement for lead creation is that it should be “as short as possible and clearly understood” – filled with information (ibid., p.176). That is why leads are difficult to write (Cappon 1982). As Schwiesau and Ohler (2013) point out, leads in online articles are observed to be shorter. The main values for lead creation are brevity and clarity.

In particular, the typical characteristics of a lead include the following aspects: First, the average lead is only 20-35 words and one to two sentences long (Bell 1991; Bednarek/Caple 2012). Second, the lead contains the summarization of the main event, including information on who, what, and where, and usually begins with who, that is, the actor of the event (Bell 1991; Cotter 2010; Bednarek/Caple 2012). Moreover, the lead can cover more than one event and provide background information or information on past events. Finally, the sources of information contained in the lead, in most cases, remain unattributed.

According to Bell (1991: 176), “the lead is the most distinctive feature of news discourse (not the headline […])”. Though the lead is located after the headline, in practice, however, the lead is usually constructed first. That is, in practice, the headline is derived from the lead (see below), which explains the high structural and functional similarity of these two elements (Kniffka 1980). For this reason, some researchers consider a headline and an intro/lead as one unit which they call the abstract (e.g. Bell 1991) or nucleus (e.g. White 1997). Conboy (2007), in turn, endorses distinguishing between headline and lead because they serve different purposes.

After headline, attribution, and lead, the actual news story – the body (in some cases, lead development) – begins. In contrast to a headline, which is, in most cases, written in the present, the news story is construed preferably in the past tense (Schudson 1995). An additional characteristic of the body is the attribution of an information source (Section 4.3.1).

Regarding the structure, a news story can consist of one or more episodes, which in turn, consist of one or more events. The elements of an event involve the actors, an action, a setting (time and place) as well as an attribution. In addition, the event may be extended by adding further elements such as a follow-up, a commentary, and a background.

Thus, the follow-up covers an action – verbal or non-verbal – which subsequently follows from the action reported of the event. The commentary, in turn, is a journalist’s or news actors’ view on the action. It can be expressed as an information of the context, e.g. by

comparing the actual action with one or more previous ones, as an evaluation of the action (positive, negative) that discusses its significance or by expressing one's expectations and making predictions on how the situation can develop next. Finally, the background refers to the situations and actions prior to the action reported. As Schwiesau and Ohler (2013) point out, the background information makes news articles more informative. Further, according to Bell (1991), follow-up and background can include episodes as well.

The development of a news story is driven by five W's and H's, namely *Who*, *When*, *Where*, *What*, *How*, *Why*, corresponding to the rhetoric of the classical antiquity *Quis? Quid? Quo? Quando? Quomodo? Cur?* which served the ancient speakers in preparing and organizing their speeches. Kniffka (1980: 200) suggests that these categories are essential for the construction of news stories. He further notes that these categories guide the mental process of news-text production and news-text comprehension. Manoff and Schudson (1987), in turn, proposed a collection of essays on the role of the W's for news production and the means for their realization.

According to e.g. Bell (1991), the modern writing of news can be characterized by a non-chronological and hierarchical order of news structural elements. The order in which events are reported in the news does not correspond to the order in which the events really happened but is rather motivated by the degree of importance of the information and its newsworthiness. Bell (1991: 172) formulated this idea as "order is everything but chronology is nothing" (with the exception of soft news). Thus, based on this, the most important and newsworthy information must be located at the beginning. Information already introduced can be repeated in the article as many times as needed, adding new, unknown details to each mention and construing a complex time structure for the article.

The idea to place the most valuable information first, which is considered as the golden principle of news article writing, has two reasons. First, the text of the news article can be easily shortened, starting from the bottom (less important information) and going to the top (more important information). Second, this kind of news article construction corresponds to the principle of selective perception. According to this principle, the news text recipient reads only what he is interested in. If the beginning of the story is not interesting, he won't continue reading and may switch to the next article.

A number of models or news schemas based on the golden principle of information organization in news articles have been developed as an orientation for news article production. These can vary for different languages. Thus, van Dijk (1988) showed that Swedish, Spanish, and Chinese follow news schemata similar to English news.

The most prominent model is that of an inverted news triangle, or an inverted pyramid, proposed in the 19th century in the USA (Bell 1991; Schwiesau/Ohler 2013, 2016;

Bednarek/Caple 2012). According to this model, the most important and newsworthy information is reported in the beginning, followed by information in descending priority. The background or contextual information is usually located at the end of the article (Bell 1991; Schwiesau/Ohler 2013; Conboy 2007; Bednarek/Caple 2012).

Another schema was proposed in White (1997). White (1997: 121) describes the structure of a news article to be an orbital structure. The headline and lead paragraph serve as the nucleus, and the following paragraphs in the body act as satellites that elaborate the nucleus by providing elaboration, appraisal, and contextualization, for example (White 1997: 121). An advantage of this model structure, as White notes, consists in the possibility of rearranging the paragraphs without disturbing the cohesion of the text (White 1997: 118).

4.3 Language of News Article

4.3.1 Reported Speech

4.3.1.1 News as an "embedded talk"

One of the particularities of news texts is the frequent usage of reported speech – direct or indirect – in the reporting on events. According to Bergler (2005), reported speech can, in some cases, make up a total of up to 90% of the news text. This observation can be explained, on one hand, by that information for news texts comes from multiple sources, or "voices", when speaking in terms of Fairclough (1995: 161), such as human (witnesses, interlocutors, etc.) or non-human (institutions, social media platforms, other news texts, official documents, etc.) (Bell 1991; Cramer 2011; Bednarek/Caple 2012). The "voices" in news texts are mediated in such a way as to construe a conversation, leading news texts to be reasonably described as an "embedded talk" (Bell 1991: 52).

On the other hand, a high preference of reported speech such as this can be seen as a result of several reasons arising out of the values which guide the process of news story production. First, reported speech serves the purpose of narrating the events (Cramer 2011: 139). In this context, it is often used to summarize an event as well as to exemplify or illustrate some point in the story. Second, reported speech enables the construction of a news story from multiple viewpoints, by this means balancing the controversial viewpoints and achieving objectivity as one of the main news values (Bednarek/Caple 2012: 91). Moreover, reported speech allows an integration of "strong evaluations" in the news story, which in another case, would violate the value of objectivity. Finally, the integration of the quotes of elite speakers can increase the newsworthiness of the reported event and the credibility of the newspaper (ibid.). Thus, the use of reported speech in the production of news texts is not an arbitrary one but rather of intentionally serving a number of purposes.

According to McDowall (1992: 137), an eyewitness is regarded to be the most preferred and the best source of information. McDowall (1992) compares the role of an eyewitness in the news article with the function of an eyewitness in court. The eyewitness citation has two more functions. First, as McDowall (1992: 136) notices, it "enables the ["your" - NKB] readers or listeners to form their own judgment of its [newspaper] credibility". Second, it allows the reporter and/or the news agency "to protect [your] company's reputation if a story is challenged" (ibid.). For selecting the eyewitness, McDowall (1992) suggests making a choice toward the authoritative source (with authority to make decisions) or an official source (often a spokesperson) that is authorized to speak on behalf of the authoritative source-eyewitness.

The principles of how information from multiple sources is integrated or mediated in news texts were the topic of many studies, and these were often discussed in connection with the concepts of intertextuality and evidentiality (e.g., Fairclough 1995; Bell 1991; Bednarek 2006; Garretson/Ädel 2008). The aspects in the focus of the studies include the ways of reporting the speech, including the types of reported speech (Section 4.3.1.2) as well as reporting expressions (Section 4.3.1.3).

Regarding a number of information sources (e.g., interview, YouTube, conversation on phone etc.) that can determine a news story, it should be noted that these can be attributed with or without naming the source, or remain unattributed. The source of information can be a person or an institution, law, etc. As Bell (1991: 190) points out, "attribution serves an important function in telling of news stories. It reminds the audience that this is an account which originated with certain persons and organizations."

Though attributed information has a higher value in the practice of news production (Bell 1991; Stenvall 2008), unattributed sources occur frequently in the news (Garretson/Ädel 2008; Stenvall 2008). If the source of information is unattributed and remains anonymous, the reasons are often provided (Stenvall 2008).

Attributed speech and mentioning a specific source can be recognized because, as mentioned in Bednarek and Caple (2012: 91), it uses linguistic devices in a way that is expected and said, without mentioning who actually alleges/expects/says.

4.3.1.2 Types of Reported Speech

Attributed material can be integrated into news articles in different ways. In this regard, Bell (1991) distinguishes between direct and indirect speech. Bednarek and Caple (2012: 92) further specify this binary classification acts in the partial (mixed) direct quote, in the free indirect speech, in the paraphrase (summary) of speech as well as in the nested (embedded) speech (reported speech in reported speech) (Table 16). The decision of which type

of reported speech is chosen, besides the structural and stylistic, also depends on pragmatic reasons (Bednarek/Caple 2012).

The use of direct speech in news texts serves three functions. First, direct speech is regarded as an "incontrovertible fact" as these are the own words of the news source (Tuchman 1980; Roeh 1982; Bell 1991). Second, the use of direct speech contributes to the distance between the journalist and the information reported, clearly distinguishing between the voice of the newspaper and the voice of the news source (Fairclough 1995). Such a distance, in turn, allows using vocabulary and phrases that would not be reproduced in indirect speech. Third, direct speech provides an authenticity, a flavor, and is "brief, pithy and colourful" in the information reported (Bell 1991: 209; Bednarek/Caple 2012: 93).

Type of reported speech	**Example**
Direct speech	"The suggestion that I was racist because of the response to Katrina represented an all-time low," Bush told a surprised Matt Auer on NBC on Monday night."
Partial direct / mixed speech, including scare quotes	He [Apari] told reporters that Berenson was "once again trying to adjust, and organizing her things" at the residence in Lima's upscale Miraflores district.
Indirect speech	He said he still felt sickened that no weapons of mass destruction were ever found in Iraq.
Free indirect speech (tense shift, use of pronouns)	Defence Minister Wayne Mapp said joint exercises were being discussed…It was also possible American soldiers would visit for joint exercises agreed between the two former allies.
Summary / paraphrase of speech act	U.S. President Barack Obama …criticized Israel on Tuesday for its decision to advance the approval of some 1,000 news housing units in East Jerusalem during a sensitive time in the peace negotiations with the Palestinians.
Nested / embedded (reported speech within reported speech)	"We really have to do something about it," Kerry said according to a Democratic official.

Table 16: Types of reported speech. *Note*: From News discourse (Bednarek/Caple 2012: 92).

However, a number of studies (e.g. Bell 1991; Garretson/Ädel 2008: 170) reveal that indirect speech is more preferred in news texts than direct speech. This can be explained by observing that indirect speech "appears to be more neutral and less immediate or vivid" (Bednarek/Caple 2012: 93). It allows distancing oneself from what is being clearly quoted, distinguishing between the voice of the newspaper and the voice of the news source (Fairclough 1995). Moreover, indirect speech "puts the journalist in control of focusing the story" (Bell 1991: 209). That is, the journalist is free to change the wordings of the speech and combine it with other information.

4.3.1.3 Reporting Expressions

A number of linguistic devices, including verbs (e.g., *says*, *believes*), nouns (e.g., *claim*, *fears that*), adverbs (e.g., *allegedly*), and prepositions (e.g., *according to*), can be used for the purpose integrating reported speech into a news article.

Bell (1991) argues that there is a relation between who is speaking and the speech verb used, the choice of which assigns a news value to the source. The older studies of Leitner (1986), Geis (1987) found out that the verb say is preferred to introduce the speech of management or the US government, and *claim* is preferred for unions.

Additionally, Bednarek/Caple (2012) point out that reporting expressions differ in the degree of their reliability. Thus, the reporting expressions such as *claim* and *allege* are less reliable than, e.g., the verb *to reveal*.

Moreover, Bell (1991), Caldas-Coulthard (1994) and Bednarek (2006), and Bednarek and Caple (2012) reveal that reporting expressions differ in respect to the functions that they fulfill. The meaning of the reporting expression is construed by the function that the reporting expressions fulfill. Due to the functions, the researchers distinguish between five types of reporting expressions, including neutral (e.g., *to say*, *to tell*, *according to*); illocutionary (e.g., *to declare*, *to announce*, *to refuse*, *to threaten*, *to insist*, *to denounce*, *to demand*, *to promise*); declarative (e.g., *to acquit*, *to plead guilty*); discourse signaling (e.g., *to add*, *to conclude*) and paralinguistic (e.g., *to whisper*, *to scream*) reporting expressions.

Thus, neutral expressions serve to integrate a reported speech, without providing or expressing any additional information included in the meaning of the verb/expression. The studies of Bednarek (2006) and Garretson and Ädel (2008) show that, based on the analysis of British and US newspapers, the usage of neutral expressions prevails in print news, to the objectivity of news reporting.

Illocutionary expressions, or news performatives (Bell 1991: 206), or illocutionary acts (Austin 1962), in turn, provide information about the purpose of the speaker. "They perform the act which they describe" (Bell 1991: 206). Performatives, according to Bell (1991: 206),

cannot be true or false. According to Fishman (1980) and Bell (1991), statements introduced by such kind of reporting expressions are newsworthy and state indisputable facts.

While declarative expressions refer to the "institutionalized linguistic act" (Bednarek/Caple 2012), discourse expressions explicitly mark the relation to the previous or the following discourse.

Finally, paralinguistic expressions provide either information on the quality of the speech or insight into someone's emotional state (e.g. fear).

4.3.2 News Actors Labeling

There are particular patterns in the labeling of news actors. One should additionally note that the news actors (who say) and the news source (which says) are often related to each other (Sigal 1987).

News actors are usually mentioned in lead and in the first paragraph (Bell 1991). The most frequent news actors in news texts, according to Bell (1991), include crime victims and criminals, politicians, celebrities, sportsmen as well as professionals. Moreover, known persons occur four times more often than unknowns (Gans 1979).

Conventionally, the news texts do not provide descriptions of the participants, which hinders the actor mentions by name and by label. The use of actor labels is, in turn, frequent. The news actors can be introduced or labeled positively or negatively. In some cases, however, the same person is labeled both positively and negatively. "It is particularly evident in situations of conflict: one side's *terrorist* is the other side's *freedom fighter*" (Bell 1991: 195). Moreover, the labels can be simple and ordinary, or they can be unique and may stick, such as the labels of Margaret Thatcher as the *Iron Lady*, and Putin as the *Alpha-Dog*.

The syntax of news actor labeling usually follows a particular scheme, serving the purpose of brevity (Bell 1991). According to this scheme, the labels are usually construed as a descriptive noun phrase and a name noun phrase (e.g., *writer-singer Salvador Sobral*, *figure-skating rising star Julia Lipnitskaya*, etc.). Moreover, they exclude any articles and prepositions (e.g., *boxing champ Mike Tyson* instead of *the boxing champion Mike Tyson*). Finally, the labels can include titles or quasi-titles of the person, these are however shortened (e.g., *Kate, Duchess of Cambridge* instead of *Kate, Her Royal Highness the Duchess of Cambridge, Countess of Strathearn, and Lady Carrickfergus*).

4.3.3 Event Categories

Cramer (2011: 76) defines event categories as "a way of naming and categorizing the events that journalists report". The most frequent realization of event categories is by means of nouns, noun phrases, and nominalizations.

Event categories indicating events can be represented in news articles by abstract nouns such as *hurricane*, *picnic*, *traffic jam*, *ball*, and *ceremony* (Bennett 1988: 14). Event categories indicating conflicting discourses are particularly typical, including nouns such as *furor*, *dustup*, *circus*, *debate*, *fight*, as well as *battle* and *war*, which indicate verbal rather than physical conflict (Cramer 2011: 76). Historical events, as Bell (1991) points out, are frequently realized as nouns combined with definite articles and attributive modifiers (e.g., *the Vietnam War*, *the Velvet Revolution*, etc.). In general, nouns are the prominent feature of news discourse (Bell 1991; Jucker 1992; Cramer 2011).

The realization of news categories as noun phrases and their role for reporting events has been studied in Aitchison (2007) and Fowler (1991).

Also, nominalizations such as *death*, *collapse*, *selection*, and *singing* can represent event categories. However, this is not a distinct criterion (Hopper/Thompson 1984: 745). Also, nouns such as *fire* and *blizzard*, which are not derived from verbs, can grammatically behave like nominalizations (Vendler 1967: 141). Nominalizations such as *information*, in turn, cannot be regarded as event categories since they lose their "eventiveness" as a result of being derived from the verb (*information* derived from *inform*) (Cramer 2011: 77).

4.3.4 Time and Place Mentions

There are patterns in mentioning time and places in news texts as well. These two essential W's (When and Where) of a news setting participate in the construction of a number of news values. Thus, time is related to the values of recency, unexpectedness, predictability, and continuity, whereas place is related to proximity and consonance (Section 4.1.3).

References to time and place in news articles can be realized in different ways. One will note that, whereas time references are in most cases simply expressed, specifying a place can be more complex, construed of multiple prepositional phrases (Bell 1991).

Thus, time and place can be anchored in news texts by means of adverbs, adverbial phrases, or adverbial clauses. The adverbials *yesterday* and *after* are, for instance, observed to be frequent in UK news discourse (Bednarek 2008). In addition, *after* and *as* are found to be favorite means for combining sequential or concurrent news events. Also, more complex adverbial constructions, including adverbial phrases such as *in the new year*, *at the moment*, *in recent weeks*, *for over twenty years*, etc., and adverbial clauses such as the one in (4.1), are often used to refer to time. Although adverbials are a preferred means for indicating time, place adverbials, such as *in our beds* in (4.1) and *clubbing* in (4.2), are also found to be frequently used to express time reference (Bednarek/Caple 2012).

(4.1) We were sleeping peacefully in our beds when the earthquake struck.

(4.2) They were both out clubbing.

In addition to adverbials, nouns are also often used to refer to a place. Place names can be used as original or derived names, adjectival forms, e.g., Russia as Soviet, etc. While referring to the same place multiple times, a variety of name variants can be such as for Russia, also Russian Federation – federation – ex-Soviet country, providing by this means an additional information and expanding or refreshing the reader's knowledge. Additionally, the place names can contain contextual or evaluative material such as *rebel Russia*.

Hallin (1987) made an interesting observation regarding place names occurring in news articles. He argues that place names often express *who*, rather than *where*. Thus, for example, the countries mentioned can perform as an actor (e.g., authority) of an event, instead of naming a geographical place or representing a set of stereotypes about a nation.

4.3.5 The Use of Numbers and Figures

Research on news language (e.g., Bell 1991; Bednarek/Caple 2012) shows that the use of numbers and figures as "the most verifiable, quantifiable, undeniable of facts" (Bell 1991: 202) is a characteristic aspect of the news text language.

Numbers and figures in news can serve two functions. On one side, they provide the reader with "objective, empirical claims" (Bell 1991: 203) and facts, achieving the objectivity and facticity of news reporting. On the other side, they serve the purpose of improving the story, or speaking in terms of van Dijk (1988: 90), "of telling good news stories", undermining the value of objectivity. Roeh and Feldman (1984: included in the title of the article) call this aspect "the rhetoric of numbers" in news. According to Roeh and Feldman (1984) as reported in Bell (1991: 203), numbers and figures are often integrated into news "for rhetorical purposes rather than to stress its facticity". The reported information in news texts, undergirded by statistics, appeals to a reader as precise and serious.

Typical news numbers, according to Bell (1991), in particular include dates, ages, number of victims or participants, weights, heights, strength, duration, etc. How often the reporter makes use of this information depends on the type of news story (Bednarek/Caple 2012). Thus, whereas the numbers and figures frequently occur in financial stories, they are relatively rare in stories about celebrities.

The use of numbers in the news, in most cases, follows common principles (McDowall 1992). That is, single digits from one to nine (sometimes ten), with exception of dates and time (e.g., *5 p.m.* instead of *five p.m.*), and digits above ten at the beginning of the sentence are written out. Sentences usually do not start with complex figures. Moreover, with the exception of national currency rates, digits are often rounded to two significant decimals and combined with expressions such as *up to*, *at least*, *more than*, *over*, and *around*. The word *billion* is usually spelled out.

4.3.6 Other Characteristics

Biber and Conrad (2009) reported a number of findings on news texts in their comparative corpus-based study of different text types.

Thus, the study reveals a rare usage of personal pronouns, in comparison to academic writing. The first- and second-person pronouns usually occur only in direct quotes. In indirect speech, they are replaced by third-person pronouns.

Also, the use of modal verbs (*may*, *can*, *will*, *should*) is less common in news texts than in academic writing. Bednarek (2008), however, reported the frequent occurrence of the modal verbs *will* and *would*, as well as of the verbs *has* and *win*, in UK news discourse.

Further, the study reports that the verbs in the present tense are observed to be slightly more frequent than the past-tense verbs. The latter is, however, more frequent in academic writing although slightly more frequent than in conversation. Additionally, Biber and Conrad (2009) observed that 15 percent of finite verbs occur in the passive, in comparison to 25 percent in academic prose, and of rare occurrence in conversation.

The occurrence of the contracted forms *don't* and *there are* not typical for news texts. However, they can occur in direct speech.

The analysis of adverbials showed that adverbials which specify place and time are very common in news texts, while linking adverbials such as *however* and *so* are less frequent. Additionally, the results indicate that a direct question rarely occurs in news texts. In case they do occur, an indirect form (e.g., *when asked about*) is preferred.

Finally, the high type-token ratio (the number of different words to the number of total words) score indicates that the vocabulary of news texts is heterogeneous, rather than repetitive.

4.4 Summary

The present chapter addressed the main characteristics of news articles focusing exclusively on news texts and leaving the use of images and videos beyond the scope.

First, after presenting the main values that guide the process of news production, the structure of news texts was addressed. It was shown that newswire texts traditionally consist of four elements, which are the headline, an attribution, an intro/lead, and the actual story – the body/lead development. The order of these elements, however, varies in the news texts and is the decision of the journalist. The body of a news article, in turn, is usually constructed non-chronologically, following the model of an inverted pyramid or triangle. According to this model, the most important, valuable, and newsworthy information is accommodated first.

Further, it was pointed out that the frequent use of reported speech is characteristic for news texts and, in some cases, can make up a total of up to 90% of the text. The source of

reported information can be a person (or institution) who (which) can be attributed with or without being named or can remain unattributed. Reported speech can be of different types, including direct speech, indirect speech, partially direct speech including scare quotes, free indirect speech, paraphrase of a speech act as well as embedded speech which is reported speech within reported speech. For the integration of reported speech, different devices can be applied, including verbs (e.g., *says*, *believes*), nouns (e.g., *claim*, *fears that*), adverbs (e.g. *allegedly*), and prepositions (e.g. *according to*).

Additionally, the frequent occurrence of news actors and event categories is characteristic for news articles. News actors can be labeled by different means such as a descriptive noun phrase and a name noun phrase (e.g., *writer-singer Salvador Sobral*, *figure-skating rising star Julia Lipnitskaya*, etc.), excluding any articles and prepositions (e.g., *boxing champ Mike Tyson* instead of *the boxing champion Mike Tyson*) as well as by shortened titles or quasi-titles of the person (e.g., *Kate, Duchess of Cambridge* instead of *Kate, Her Royal Highness the Duchess of Cambridge, Countess of Strathearn, and Lady Carrickfergus*). The most frequent realization of event categories is by means of nouns, noun phrases, and nominalizations.

Moreover, according to the addressed studies, news texts, in general, are characterized by the use of a standard language and the rare occurrence of metaphors. Further, it was reported that adverbials which specify place and time are very common in news texts while linking adverbials such as *however* and *so* are less frequent. Also the use of modal verbs (*may*, *can*, *will*, *should*) is less common in news texts than e.g. in academic writing. In turn, present tense verbs are slightly more frequent than past tense verbs. Further, the occurrence of the contracted forms *don't* and of *there are* is not typical for news texts. The contracted forms can, however, occur in direct speech. An occurrence of a direct question is rare as well. Finally, another characteristic of news texts is the frequent use of numbers and figures.

5 Typology Construction: Types of Contradictions in News Texts

Chapter 5 describes the methodology and reports the results of a corpus-based typology construction of contradictions occurring in news texts. Two dimensions for typology construction are in the focus here: contradiction cues which can be used for processing by a machine (Section 5.2.1) and meaning similarity or relatedness of the contradiction parts (Section 5.2.2). The chapter is organized as follows: First, Section 5.1 outlines a two-stage process of contradictions corpus compilation, including data collection (Section 5.1.1) and data validation and filtering (Section 5.1.2). Based on the results of this section, Section 5.2 summarizes the process of typology construction and provides the description of the identified types. The validation of the constructed typology, the process of corpus annotation, as well as the description of the resulting Gießen Annotated Corpus of News Contradictions, are the topics of Section 5.3.

5.1 Compilation of News Text Contradiction Corpus

5.1.1 Data Collection

5.1.1.1 Collection of News Texts

The previous studies on contradictions emphasized mainly the manual construction of contradictions by human participants. Different methods have been proposed for this purpose (see Section 2.2). Though by means of manual construction, a large number of contradictions can be produced in an adequate amount of time, the contradictions constructed may not represent the whole spectrum of the naturally occurring contradictions. To avoid this problem, contradictions for the purpose of this study were collected by human participants reading sets of news texts on a common topic. However, the selected method is characterized by a high complexity due to the recognition of contradictions by humans, especially of implicit ones (see Section 3.4.1 for the definition of implicit contradictions), is a challenging task which, besides world knowledge (personal and cultural knowledge), requires an analytical ability.

For news text collection, published online news stories by world news agencies as original source of the news stories were preferred. In total, 15 agencies from nine countries (China, France, United Kingdom, Italy, Russia, South Korea, Qatar, Ukraine, U.S.A.) were selected. In addition to the news agencies, the New York Times as one of the prominent newspapers in the USA and in the world, as well as the NBC News as an online performance of the NBC – National Broadcasting Company of the U.S., were preferred as a data source as well. Due to the monolinguistic orientation of the study, the English portrayal of the chosen sources is considered. That is, only news texts published in English were obtained. The distribution of news articles according to the topic of the news story, date of publishing, and their source is presented in Table 17.

As can be seen in the Table 17, the text data obtained represents either soft or hard news (for definition of soft and hard news, see Section 4.1.1) on a variety world events such as an armed invasion of Russia in Crimea (February-March 2014), a fire in a Brazil nightclub (January 2013), first protests after the killing of a black teen in Ferguson (August 2014), the mysterious vanishing of the Malaysia Airlines Flight MH370 plane (March 2014), the shooting down of the Malaysia Airlines Flight MH17 (July 2014), the military coup in Thailand (May 2014), the natural catastrophe in the Philippines (November 2013), the crash of the

Topic / News Agency or Newspaper (Country)	**Date**	**Agence France-Presse (France)**	**AL JAZEERA (Qatar)**	**ANSA (Italy)**	**BBC (UK)**	**BLOOMBERG (USA)**	**China News Service (China)**	**CNN (USA)**	**ITAR-TASS (Russia)**	**NBC News (USA)**	**REUTERS (UK)**	**RIA Novosti (Russia)**	**The New York Times (USA)**	**UNIAN (Ukraine)**	**XINHUA (China)**	**Yonhap News Agency (South Korea)**
Armed Invasion in Crimea	26.02.14 – 01.03.14	2	3	-	2	2	-	-	2	1	2	2	1	2	2	-
Brazil Nightclub Fire	27.01.13 – 29.01.13	-	1	-	4	3	1	2	1	-	-	1	2	-	-	-
Costa Concordia Disaster	13.01.12 – 17.01.12	-	-	-	3	-	-	4	1	-	3	-	1	-	1	1
Death of a Black Teen in Ferguson – First Protests	10.08.14 – 14.08.14	2	2	-	2	1	-	2	-	4	3	-	1	-	-	-
Malaysia Airlines Flight MH370 Missing	07.03.14 – 10.03.14	-	2	-	2	2	2	2	1	-	2	1	2	-	4	-
Malaysia Airlines Flight MH370 Shot Down	17.07.14 – 20.07.14	-	2	2	2	2	3	-	2	-	2	4	3	2	2	1
Military Coup in Thailand	22.05.14 – 24.05.14	3	2	-	2	2	-	2	2	2	2	2	-	-	2	-
Typhoon Haiyan Hits Philippine	07.11.13 – 10.11.13	-	2	-	2	3	-	1	-	-	2	-	2	-	2	-
Whitney Houston's Death	11.02.12 – 14.02.12	-	2	-	2	2	-	3	-	-	2	1	2	-	2	-
Total																**165**

Table 17: The distribution of news articles according to the topic of the news story, their source, and date of publishing.

Costa Concordia ship (January 2012) as well as the unexpected death of Whitney Houston (February 2012). All of these events and, in particular, the invasion of Russia into Crimea, the shot-down Malaysian plane, and the death of the black teen in Ferguson evoked a high resonance in the world, which resulted in numerous controversy discussions, especially in the social media. The news articles, in turn, have been characterized by a number of conflicting opinions and divergent facts. Other events, such as the natural catastrophe in Philippine, the disaster with the Costa Concordia ship, and the death of Whitney Houston, were reported in the media, delivering a number of contradictory facts as well.

The news articles for the corpus were obtained randomly. Only the news articles were considered that appeared within three days after the actual event had happened, assuming that such news stories most probably include contradictions. Additionally, the focus on three to four days covers the time difference between the countries that are located in different time zones (e.g. the time difference between Moscow (Russia) and New York (USA) is 7 hours). Finally, it should be noted that only original articles without any later corrections were considered. In total, 165 news articles were collected. Due to their irrelevance for the purpose of this study, images and videos have been removed from the files, leaving only the plain news texts.

5.1.1.2 Survey 1: Finding the Contradictions

To manually find the contradictions in the 165 news texts collected, for the reason of convenient processing, the news texts were grouped into 15 questionnaires with two topics per questionnaire. Each questionnaire consists of one page with instructions for conducting the survey, one page with a definition, and some examples of contradiction as well as news texts on two world events (henceforth topic), including two forms for filling in – one for each topic – containing a column for the IDs of the contradictory sentences and a column that explains the decision made (an example of a questionnaire can be found in Attachment A). Each topic contains about four to eight news texts of about 120 sentences, which amounts to approximately 17,000 tokens per topic. Each news text was provided with metadata (e.g., the title of the article, the date of publishing) and a plain text. The sentences of the texts with a common topic were subsequently enumerated. As mentioned above, due to the purpose of the study defined, video and image materials were excluded from the survey.

There was a total of 30 participants in the survey, mainly first term Master students of computational linguistics, English and German linguistics of the University of Giessen. Table 18 presents the distribution of the English language competencies among the participants. The share of the knowledge level at English as a native language to good knowledge amounts to 80% in total.

Level of Competency in English	Number of students	%
Native speaker	1	3.3
Fluent	10	33.3
Very good	5	16.7
Good	8	26.7
Basic knowledge	4	13.3
Unknown	2	6.7

Table 18: Distribution of English language competencies.

At the beginning of the survey, the participants were provided with the questionnaires and were informed about the possible occurrence of contradictions in the news texts. According to findings reported by Winograd and Johnston (1982), Glenberg et al. (1982), August et al. (1984) as well as Baker and Zimlin (1989), a prior notification for a reader about any contradictions occurring in a text makes the reader process the text more concentratedly and by this, increases the probability that she/he would recognize contradictions occurring in text. In addition, the participants were asked to process the survey questionnaires at home and were given a time limit of one month. Moreover, they were allowed to use all kinds of supplementary material, such as encyclopedias, dictionaries, calculators, etc., without any limitations.

5.1.1.3 Results and Evaluation

In total, 1 out of the 30 questionnaires was returned completely not filled in and was thus considered invalid for the study. Further, 4 out of 30 questionnaires (5, 8, 8K, 10) were only partially filled in, i.e. the texts concerning only one topic had been processed. In this case, only that part of the questionnaires was taken into consideration for compiling the corpus.

The fully or partially processed questionnaires can be explained by at least three reasons: 1) there were no contradictions occurred in news texts, 2) the survey participant was not able to find any contradictions as a result of, e.g., insufficient English knowledge or/and world knowledge, 3) the survey participant was not willing to process the questionnaire.

In total, 943 pairs of (potentially) contradictory or contrary sentences were collected during this stage of Survey 1. As some questionnaires were not properly processed according to the instructions provided, the step of data standardization was additionally applied. That is, e.g. in case the participants had listed the contradictory sentences in a line, these were converted to a pairwise presentation.

In a closer analysis of the data, it could be observed that a number of sentence pairs, half of all the cases, had no clear contradictions (519 instances out of 943). Either both parts of the sentence pairs were equivalent in expressing the same proposition (151 instances), or

they were related but without exhibiting a contradictory or contrary element (200 instances), or they were absolutely unrelated (94 instances). Additionally, in the case of 14 pairs, some of the sentences mentioned in the pair did not exist. Finally, 58 instances were identified as duplicates. One of the possible reasons for these findings is that the survey instructions were violated. That is, as earlier mentioned, the sentences recognized as contradictory that were listed in a line in the questionnaire were not all contradictory, which could be detected after converting them into a pairwise presentation. An inadequate definition of contradiction that was provided in the questionnaires can be considered as another reason as well.

Because the present study is computationally oriented, the quality of the data has an impact on the development and evaluation of the system for detecting contradictions. For this reason, an additional survey was conducted in order to validate and filter the contradictions collected. Cases which were clearly identified as non-contradictions were excluded from Survey 2. Thus, only 424 pairs of (potentially) contradictory sentence pairs in the survey were validated further. The description of Survey 2, including results, is provided in the next Section 5.1.2.

5.1.2 Data Validation and Filtering

5.1.2.1 Survey 2: Contradiction or Not

A second survey was conducted to validate the 424 (potential) contradictions collected. The participants of the study had to decide whether each of the given pairs of sentences is contradictory or not. The agreement of the raters on contradictions has been then evaluated by computing an inter-agreement score.

At the beginning of Survey 2, the participants were provided with a questionnaire containing 424 sentence pairs along with the original news texts in order to clarify the unclear time reference of the statements in the sentence pairs when required. Moreover, in contrast to Survey 1, in addition to a definition of contradiction and examples, the participants were provided with a Test for Contradiction (Appendix B) to increase the number of correct decisions.

Three graduates of economics from different German universities took part in the survey. Regarding the argumentation on choosing the number of participants, see e.g. Snow (2008). All three participants could attest a fluent level of English language competency. For the processing of the study, the participants (henceforth raters) were given one month and were allowed to use all supplementary material needed (printed and online encyclopedia, dictionaries, etc.) without limitations.

5.1.2.2 Results and Evaluation

The distribution of agreement and disagreement on contradictions among the three survey participants and the discussion of the given 424 sentence pairs before (second column) and after (third column) is summarized in Table 19.

Decisions of raters on contradiction	Number of sentence pairs before discussion	Number of sentence pairs after discussion
3 (yes)	250	311
1 (yes) / 2 (no)	13	0
1 (no) / 2 (yes)	59	0
3 (no)	102	113
Total		**424**

Table 19: Distribution of agreement and disagreement on contradictions among the raters.

As can be obtained from the table, all three raters agreed on 352 cases (250 unanimously positive decisions and 102 negative decisions) out of 424 before the discussion. For of 72 (1 yes/2 no; 1 no/2 yes) of the 424 sentence pairs, only two raters had made the same decision. The overall agreement among the raters and reliability of their decisions before the discussion is given in Table 20. For computing the inter-rater agreement, the best methods for working with the nominal data of multiple raters are those by Fleiss's Kappa (Fleiss 1971) and Krippendorff's Alpha (Krippendorff 2012) (for more information on computing the agreement between multiple raters with nominal data, see Scott 1955; Cohen 1960; Light 1971; Conger 1980; Davies/Fleiss 1982; Carletta 1996). For computing the scores, the online tool developed by J. Geertzen[20] was applied.

	Fleiss			**Krippendorff**		
	A_obs	A_exp	Kappa	D_obs	D_exp	Alpha
Score before discussion	0.887	0.574	0.734	0.113	0.426	0.734
Score after discussion	1.0	0.609	1.0	0	0.391	1.0

Table 20: The scores of inter-rater agreements on contradictions before and after discussion, computed with Fleiss's Kappa and Krippendorff's Alpha.

In total, 424 cases with 1272 decisions of three raters have been analyzed. There were no missing values in the data. The results of the survey before the discussion already indicate a high overlap of rater decisions, which amounts to 0.734 (73%) for both Fleiss's Kappa and Krippendorff's Alpha.

[20] https://nlp-ml.io/jg/software/ira/

From the perspective of corpus compilation, the cases where only one rater disagreed are of most interest for us as these can be clarified and possibly resolved, to increase by this means the size of the dataset for development and evaluation of the system.

For this reason, an additional discussion between the raters on their decisions for 72 cases took place. The following reasons for disagreement on contradictions/contrarieties during the discussion could be identified:

- The sentence pairs were mistakenly identified as contradictions/contrarieties,
- The rater did not notice contradictions/contrarieties,
- The raters operated with different interpretations of measures, factuality of sentences, word meanings,
- The raters made wrong inferences,
- Missing world knowledge on the required topic.

During the discussion, a consensus between the raters could be achieved for all 72 cases. The distribution of the decisions after the discussion is given in Table 19. As can be seen in the table, the number of recognized contradictions for the dataset could be increased from 250 to 311 cases. The Fleiss's and Krippendorff's scores computed after the discussion are provided in Table 20, indicating as expected and proving a full agreement among the raters. Both scores amount to 1.0 (100%).

Thus, as the result of Survey 1, Survey 2, and Survey 3, a corpus of contradictions including 311 instances could be compiled. This corpus served as the data basis for constructing a typology of contradictions as well as for developing and evaluating a CD system (Chapter 7). The results of the typology construction are reported in the following Section 5.2.

5.2 Typology of Contradictions

5.2.1 Dimension: Contradiction Cues

The typology we propose in the present section has been developed from the perspective of computational processing, aiming at the identifying the linguistic and non-linguistic features that can serve as cues for an automatic CD task. It should be mentioned that the typology developed can be a matter of controversy when trying to justify it through the prism of contradictions processing by humans, especially in case of explicit and implicit contradictions (see below).

In general, based on the commonalities between corpus instances with respect to the way of their realization or cues, we distinguish between the following kinds of contradictions occurring in news texts according to their contradictory cue: negation, opposition, numerical, lexical, and factual (described in the present subsection). These kinds can be realized

either explicitly or implicitly (Section 5.2.2), constructing by this means the following types of contradictions: explicit-negation, explicit-opposition, explicit-numerical, explicit-lexical, implicit-negation, implicit-opposition, implicit-numerical, implicit-lexical, and implicit-factual. It is noted that since world knowledge is required to recognize factual contradiction, all instances of factual contradictions are regarded as implicit. The nine types of contradictions mentioned are to be regarded as only hypotheses for now. The validation process of the typology developed will be described in Section 5.3.1. The construction of typology has been conducted manually.

The first kind of contradictions that we would like to address according to contradiction cue, is that arising from the occurrence of negation. That is, one part of contradiction negates what is said in the second part. We refer to this kind of contradiction as negation. Following Aristotle's formulation of the LNC (Section 3.1.1), contradiction arising from negation can be regarded as prototypical. However, as shown in Section 3.2.1, expression of negation in natural languages is not limited to the simple negation operator *not* and occurs in a variety of realization forms. Also, the aspect of multiple negation has to be mentioned in this concern (Section 3.2.4).

The realization mechanisms of negation contradictions are represented by a number of features. Thus, a simple indicator of something which is asserted in one part of contradiction that is negated in the second one is the presence of a syntactic negation such as the negative particles *not* and *n't*, the determiner no, and the pronouns *no one*, *nothing*, *nobody*, *never*, and *nowhere* (5.1).[21]

(5.1) Ukraine's Defense Ministry said on its website late yesterday that the army *hadn't used missiles* in its operation against the separatists.

Russia's Defense Ministry said Ukraine *had been using a tracking station linked to a Buk-M missile system* near Donetsk yesterday, according to the RIA Novosti news service. (ID 229)

Moreover, the features *not*, *n't*, and *no* frequently occur in news texts in combination with nominalizations such as, e.g., *no evidence*/*signs*/*markings*/*indication*/*confirmation of/that*, *no information on/about* + *any* + *subject*, *no signs of* + *any* + *subject*, *don't have information on* + *subject*, etc. as exemplified in (5.2). Such kinds of constructions are often combined with inherent negatives such as *to fail to find evidence*.

(5.2) Al Jazeera's Rob McBride, reporting from Beijing, said that it is a very hard situation for the airline as it *does not have the visual confirmation that its plane has crashed*.

[21] All examples of contradictions provided in this and the following sections are supplied with an ID as defined in the final corpus (Section 5.3).

> *A Malaysia Airlines flight* carrying 227 passengers and 12 crew *crashed* into the sea 153 miles off the coast of Vietnam's Tho Chu island on Saturday, according to a Vietnamese navy officer quoted by state media. (ID 001)

Beside syntactic, also morphological negation (e.g., *unidentified*, *unusual*, *illegally*) such as in (5.3) can indicate a negation contradiction.

(5.3) There were scenes of panic as the Costa Concordia *hit a sandbar* on Friday evening near the island of Giglio and listed about 20 degrees.

> The Costa Concordia, carrying 4,200 passengers and crew on a weeklong Mediterranean cruise, *slammed into an undetermined object* near the island of Giglio on Friday night as passengers for the late seating had just started dinner, tucking into appetizers of grilled mushrooms and scallops. (ID 283)

A special case of negation contradictions is represented by contradictions realized by using inherent negation – words of positive form but of negative meaning (*to fail*, *to deny*, *to close*, *to stop*, *a problem* etc.) (5.4).

(5.4) Simferopol International Airport currently works in its normal mode, *the flights are performed regularly*.

> They have occupied runway and *all plane movements have been stopped*, the news agency quoted the source as saying. (ID 079)

Further analysis is required when inherent negatives occur in both parts of a contradiction and in equivalent positions, as in the case of *crashed* – *shot down* in the example (6.9b) below. In this case, in order that a contradiction can be detected by a machine, the processing of the lexical meaning is additionally needed (see below for a lexical type of contradiction).

The next kind of contradictions, which we call opposition, arises in the case of opposition meaning relations between the sentence parts of contradiction such as exemplified in (5.5). The opposition meaning relation is observed here between *bodyguard* – *hairdresser*.

(5.5) Authorities have said that police and fire officials were called to Houston's room at the Beverly Hilton Hotel at 3:43 p.m. Saturday after *her bodyguard* found her unconscious body in a bathtub.

> The singer *was* found unconscious and submerged in the bathtub of her room at the Beverly Hilton by *her hairdresser* Saturday afternoon, according to TMZ. (ID 201)

Opposition meaning relations consist of antonyms, directional opposites, complementaries, converses (failed converses in cases of contradictions), and heteronyms. Cues of opposition contradiction also include to some degree meronyms of the same object. Löbner (2013: 208-220) defines the listed meaning relations as follows:

- Antonym: Two expressions are antonyms iff they express two opposite extremes out of a range of possibilities (e.g., *old* – *young*, *light* – *dark* etc.). Antonyms are logically incompatible, but not complementary;

- Directional opposites: Two expressions are directional opposites iff they express opposite cases with respect to a common axis (e.g., *to begin – to stop*, *to appear – to disappear*, *yesterday – tomorrow* etc.);
- Complementaries: Two expressions are complementaries iff they express an either-or alternative in some domain (e.g., *free – occupied*, *to buy – to rent*, *female – male* etc.);
- Converses: Two expressions are converses of each other iff they express the same relation between two entities, but with reversed roles (e.g., *x is above y – y is below x*);
- Heteronyms: Expressions are heteronyms iff they denote alternatives in some domain of more than two possibilities (e.g., *color terms*, *numbers words* etc.);
- Mereology, meronymy, and holonymy: A set of expressions forms a hierarchy in terms of holonyms and meronyms, where A is a meronym of B, and B a holonym of A, iff A denotes constitutive parts of the kind of things that B denotes (e.g., *head* and *neck* are meronyms of *body* and *body* is holonym for *head* and *neck*).

We define numerical contradictions as all contradictions that arise as the result of divergent or incompatible expressions of time and date, quantity, and quality (e.g. weight) such as exemplified in (5.6). Contradictions of this type give incompatible answers to the questions: How much/many? Which one? Numbers and numerals belong to this kind of cues. We treat numerals in a broad sense, referring to all expressions which deal with numbers. These include cardinal numbers (*one*, *two*, *three*, etc.), specialized numbers (*zero*, *couple*, *oh*, *deuce*, *trio*, *solo*, *sextet*, *dozen*, etc.), adverbs of frequency (*once*, *twice*, etc.), ordinal numbers (*first*, *second*, *third*, etc.) as well as fractions and decimals (e.g., *one sixteenths*, *one third*, etc.). Also, quantifiers (e.g. *all*) are regarded as cues of numerical contradictions. Furthermore, in the case that the negative quantifiers, *no one* and *nobody*, are used in combination with numbers or numerals, we speak of a numerical kind of contradiction.

(5.6) The passengers were of *14 different nationalities*, Mr. Jauhari said.

The passengers are of *13 nationalities*, the airline said. (ID 22)

It is to be noted that numerical contradiction cues can occur in different combinations with each other (e.g., as a word in one part of a contradiction and as a number in the second part). Moreover, for an efficient automatic processing, in most cases, the numbers and numerals combined with diverse metrics (e.g., liter, kg) have to be converted to a common comparable basis such as, e.g. in case of 2 kg and 3000 g. Furthermore, numbers and numerals can be specified by means of adverbs of degree, such as *about*, *approximately*, *at least*, *more than*, *up to*, and *less than* (5.7), which makes automatic CD challenging (implicit contradiction, see next section). Also, performing logical operations such as addition,

multiplication, etc. (5.8) can be of importance in detecting the (implicit) divergent use of numbers and numerals. Finally, when correctly identifying whether divergent time mentions are contradictory or not, different time zones have to be taken into consideration as well in analyzing the sentence pairs. For these reasons, contradictions should not be regarded as isolated from the text they occur in.

(5.7) With *sustained winds of 315 kph (195 mph)* and gusts as strong as 380 kph (235 mph), Haiyan was probably the strongest tropical cyclone to hit land anywhere in the world in recorded history.

Haiyan had *top winds of almost 196 miles (315 kilometers) per hour* when it was about 489 miles southeast of Manila, the U.S. Navy's Joint Typhoon Warning Center said at 2 p.m. East Coast time. (ID 061)

(5.8) The plane crashed on Thursday, killing all *298 people on board*.

A Malaysian flight crashed Thursday in eastern Ukraine near the Russian border, with all the *280 passengers and 15 crew members on board* reportedly having been killed. (ID 227)

Another kind of contradiction – the lexical one – arises as the result of using words with incompatible meanings. Thus, in example (5.9 a and b), the contradiction can be detected only when knowing the exact meaning of the concepts/words *fresher's ball* and *summer break*, and *to shot down* and *to crash*, respectively.

(5.9) a. According to Diario de Santa Maria, students from the city's federal university (UFSM) were holding a *freshers' ball*.

The club was hosting a party for a group of students from the Federal University of Santa Maria who were celebrating the end of their *summer break*. (ID 178)

b. The plane *crashed* on Thursday, killing all 298 people on board.

Malaysia Airlines flight MH17, while en route from Amsterdam to Kuala Lumpur, was reportedly *shot down* by a surface-to-air missile on Thursday in war-torn Ukraine, killing all 298 people on board. (ID 228)

Further, the simultaneous usage of incompatible concepts as in (6.10) signals a lexical contradiction. Mučnik (1985) refers to this kind of contradiction as an abstract (see Section 3.4.2 for the definition).

(5.10) All you could *see* was *noise and creaks*. (ID 308)

Contradictions of a lexical kind also appear as the result of differences in the expression of possibility, probability or necessity toward the real world based on the knowledge of facts as in (6.11).

(5.11) A Malaysia Airlines flight carrying 227 passengers and 12 crew has gone missing over the South China Sea, *presumed crashed*.

A Malaysia Airlines flight carrying 227 passengers and 12 crew *crashed* into the sea 153 miles off the coast of Vietnam's Tho Chu island on Saturday, according to a Vietnamese navy officer quoted by state media. (ID 000)

That is, by saying that a plane is presumed crashed, the speaker, based on his knowledge of the facts and previous experience, believes that the flight has crashed but is not certain about this. The utterance expresses the possibility that a plane has crashed but does not state it as an established/confirmed fact. Whereas, when saying that a plane crashed, the speaker states that the crash of the plane is indeed confirmed as fact. The two statements in (5.11), if referring to the same plane and being uttered at the same time, are thus contradictory.

Also, we regard the tautological use of terms – immediate unnecessary repetition of a word or concept with the same meaning – as a cue of lexical contradictions (5.12). Though we deal with a tautology here, it is often cognitively perceived by a human as a contradiction (a psychological contradiction in terms of Mučnik 1985).

(5.12) "I would also like to say to the *Brazilian people* and to the *people of Santa Maria* that we stand together at this time, and that even though there's a lot of sadness, we will pull through," she said, speaking from Chile. (ID 181)

Finally, besides the above-listed cues of lexical contradiction, also conjunctive adverbs, such as *however*, *but*, and *in contrast* (5.13), which show that a contrast relation between clauses can be a strong indicator of an occurring contradiction. As can be observed in the corpus, contradictions arising from the use of conjunctive adverbs are contact contradictions.

(5.13) Police said Brown was shot in a struggle with a gun in the police car but have not said why Brown was in the car.

But a witness to the shooting interviewed on local media has said that Brown had been putting his hands up to surrender when he was killed. (ID 145)

We want to address the next type of contradiction according to a contradiction cue, those that require knowledge of the facts in order to be detected (5.14). We refer to this kind of contradiction as factual.

(5.14) By Saturday, the storm had left the Philippines, on a path to Vietnam, according to *the Joint Typhoon Warning Center in Honolulu*. (ID 071)

In order to identify the contradiction in (5.14), the knowledge that the Joint Typhoon Warning Center is located in Pearl Harbor (Hawaii), and not in Honolulu (Hawaii), is essential.

Another cue of factual contradiction is represented by the wrong or divergent spelling of the same name referring to the same individual, institution, etc., as exemplified in (5.15) as the result of a typo or missing knowledge. This kind of cue can be detected by consulting an additional knowledge source, in case of multiple sentences in which there are different spellings. The main difficulty hereby is to determine whether the differently spelled names refer to the same object in the world. Only co-referent, differently written expressions can be judged as contradictions.

(5.15) The pilot was *Capt Zaharie Ahmad Shah*, 53, who joined Malaysia Airlines in 1981, Mr. Yahya said.

The flight was piloted by *Captain Zahaire Ahmad Shah*, a Malaysian aged 53. (ID 009

5.2.2 Dimension: Relatedness of the Parts

One of the main difficulties of finding contradictions by machine, before the actual identification of contradiction cue can take place, is to find two or more sentences that are related to each other and convey contradictory information on the same concern. First, this problem arises when parts of a contradiction are often separated by time in a portion of a text or even by whole texts (distant contradictions vs. contact contradictions). Second, the parts of contradictions, in most cases, are not verbally or syntactically equivalent or similar (implicit contradictions vs. explicit contradictions) as discussed in Svintsov (1979) (see Section 3.4.1). Svintsov, however, does not provide further details on how implicit contradictions can be realized and limits his discussion to single examples. Taking this problem into consideration, in turn, we made an attempt to describe explicitly the implicit contradictions, based on the examples of contradictions from the compiled corpus of news text contradictions.

Following Svintsov (1979), we define a contradiction as explicit when its parts or elements of these parts are verbally (lexically) and structurally equivalent or similar. We consider the parts (or their elements) of a contradiction as verbally similar in the case of using paraphrasing at word level realized by synonyms, antonyms along with negation, converse substitution, change of voice, change of person, pronoun/co-referent substitution, repetition/ellipsis, function words variations, actor/action substitution, verb substitution, manipulator/device substitution, general/specific substitution, metaphor substitution, part/whole substitution, verb/noun conversion, verb/adjective substitution, verb/adverb conversion, noun/adjective conversion, verb-preposition/noun substitution, change of tense, change of aspect, change of modality, semantic implication, approximate numerical equivalence and external knowledge as summarized in Bhagat and Hovy (2013) based on theories of Mel'čuk (2012), Honeck (1971) and Harris (1981).

In general, Bhagat and Hovy (2013: 1) define paraphrase as "sentences or phrases that convey the same meaning using different wording". According to de Beaugrande and Dressler (1981: 49) paraphrase refers to the process of "repeating content but conveying it with different expressions". In her turn, Nevěřilová (2014: 556) states that "paraphrase *s'* of a sentence *s* is a sentence that has the same or almost the same meaning as *s* in a given context" and can be regarded as a mutual entailment (addressed below), that is $s \rightarrow s'$ and $s' \rightarrow s$. Paraphrases can be realized on different levels of granularity such as at word, sentence, paragraph or document (text) level and be of different nature. In the present study,

paraphrases at word and sentence level are considered. Further, according to the nature of paraphrase, one distinguishes between lexical paraphrases realized by the use of different words, syntactical paraphrases which are the result of using different structures, referential and idiomatic.

In contrast to the explicit realization of contradictions, the parts of implicit contradictions exhibit little or no verbal or structural equivalence. For their detection by a system, a deeper meaning analysis, inferencing, and computational-adapted sources of world knowledge are required. Thus, in the previous section, it was already mentioned that the detection of implicit-numerical contradictions requires performing mathematical operations (6.8) or/and analyzing the adverbs of degree which often come along with numbers and numerals (6.7). Moreover, we count all factual contradictions as implicit contradictions because additional inferencing and/or specific world knowledge is needed. Finally, we regard as implicit the contradictions related to the parts realized by paraphrasing at sentence level.

Since Frege (1892/2011), it is scientifically accepted that a proposition has at least two levels of propositional content. The contemporary semantics distinguish between three levels of meaning, including expression, utterance, and communicative levels (Löbner 2013). While expression meaning is merely based on the linguistic material of the sentence, utterance meaning results from expression meaning used and interpreted in a certain context by fixing the references. Communicative meaning, in turn, is the meaning of an utterance in a given social interaction. Therefore, "what a speaker intends to communicate is characteristically far richer than what she directly expresses; linguistic meaning radically underdetermines the message conveyed and understood" (Horn 2004: 3). The sentences convey more information than they express verbally or explicitly. Implicitly encoded information is referred to as implication (e.g., Nørgård-Sørensen 1992; Averintseva-Klisch 2013). In turn, the process of determining or decoding the implicit meaning of sentences, as well as implicit connections between the sentences, is termed as inference (Beaugrande-Dressler 1981; Hobbs 1983; Velde 1989; Brown/Yule 1991).

Inferences can be of different types. On one side, people can make use of logical inferencing via deduction (way of reasoning from general premises to a particular conclusion), such as in chess example (3.36) discussed in Section 3.4.1, but also via induction (way of reasoning from particular premises to general conclusions) and abduction (inference to the best explanation). However, as Brown and Yule (1991: 34) point out, logical inferencing is rarely used in everyday discourse as people tend to "operate with a rather loose form of inferencing" as exemplified in (5.16), "which have some likelihood of being justified". Brown and Yule term this kind of inferencing as pragmatic inference.

(5.16) a. John was on his way to school. (Sanford/Garrod 1981: 10 cited in Brown/Yule 1991: 34)

b. Last week he had been unable to control the class.

By reading (5.16a) first, the most readers will infer that John is a schoolboy. However, by further reading (5.16b) – a sentence which follows (5.16a) in the same text – the reader will obviously give up his original inference and make a new one, namely that John is a schoolteacher. To be able to make such kind of inferences, the reader makes use of his sociocultural knowledge.

In general, scientists distinguish between two kinds of pragmatic inferences. On one side, there are the pragmatic inferences that are closely bound to the words and syntax of the text. This is referred to as text-bound (Graesser, Singer and Trabasso 1994; Perfetti/Stafura 2015), encoded (Kintsch 1988; van den Broek 1994; Singer et al. 1994), or language-based inferencing (Averintseva-Klisch 2013). The basis for text-bound inferencing is represented, according to Yule and Brown (1991), by implications such as pragmatic presupposition as well as conventional and conversational implicature. In addition to these types of implication, Nørgård-Sørensen (1992), in turn, regards logical entailment, logical (semantic) presupposition, and conventional and conversational implicature as the carriers of implicit information. He further terms logical entailment and logical presupposition as logical implications and two kinds of implicature as a pragmatic implication. Logical implications are defined in terms of truth values and are "a feature of a sentence" (Nørgård-Sørensen 1992: 27), in contrast to pragmatic implication which is a part of utterance meaning. Finally, implications addressed by Averintseva-Klisch (2013) include logical presupposition and conversational implicature. All types of implication mentioned will be described below.

On the other side, besides text-bound, the scientists also distinguish knowledge-driven or knowledge-based (Graesser et al. 1994; Averintseva-Klisch 2013; Perfetti/Stafura 2015) or activated (Kintsch 1988; van den Broek 1994; Singer et al. 1994) pragmatic inferences which are constructed based on the individual world knowledge of the reader as well as on the world knowledge of a certain cultural group. Knowledge-driven inferences are independent of the text surface. Examples of knowledge-based pragmatic inferences from a parable by Ambrose Bierce "How Leisure Came" have been proposed in e.g. Graesser et al. (1994) and are presented in Table 21. The text of the parable is as follows:

> A Man to Whom Time was Money, and who was bolting his breakfast in order to catch a train, had leaned his newspaper against the sugar bowl and was reading as he ate. In his haste and abstraction, he stuck a pickle fork into his right eye, and on removing the fork the eye came with it. In buying spectacles the needless outlay for the right lens soon reduced him to poverty, and the Man to Whom Time Was Money had to sustain life by fishing from the end of the wharf.

Type of Inference	Brief description	Text that elicits Inference	Inferences
Referential	A word or phrase Is referentially tied to a previous element or constituent in the text (explicit or inferred).	"...on removing the fork the eye came with it"	Fork is the referent for it.
Case structure role assignment	An explicit noun phrase is assigned to a particular case structure role, e.g., agent, recipient, object, location, time.	"the man leaned his newspaper against the sugarbowl"	Against the sugarbowl is assigned to a location role.
Causal antecedent	The inference is on a causal chain (bridge) between the current explicit action, event, or state and the previous passage context.	"In his haste and abstraction he stuck a pickle fork into his right eye..."	The man was careless and mis-aimed his fork.
Superordinate goal	The inference is a goal that motivates an agent's intentional action.	"A Man to Whom Time was Money, and who was bolting his breakfast in order to catch a train..."	The man wanted to get to work and earn money.
Thematic	This is a main point or moral of the text.	The entire passage	Haste makes waste.
Character emotional reaction	The inference is an emotion experienced by a character, caused by or in response to an event or action.	"...the needless outlay reduced him to poverty"	The man became sad.
Causal consequence	The inference is on a forecasted causal chain, including physical events and new plans of agents.	"...on removing the fork the eye came with it"	The man became blind in his right eye.
Instantiation of noun category	The inference is a subcategory or a particular exemplar that instantiates an explicit noun or an implicit case role that is required by the verb.	"...breakfast..."	Bacon and eggs.
Instrument	The inference is an object, part of the body, or resource used when an agent executes an intentional action.	“...the Man to Whom Time was Money had to sustain life by fishing from the end of a wharf”	The man used a rod and reel (to fish).
Subordinate goal-action	The inference is a goal, plan, or action that specifies how an agent's action is achieved.	“...who was bolting his breakfast"	The man grasped his fork and moved it toward his mouth.
State	The inference is an ongoing state, from	“...the Man to Whom Time was Money had to sustain	Fishermen are poor; the city has a wharf.

	the time frame of the text, that is not causally related to the story plot. The states include an agent's traits, knowledge, and beliefs; the properties of objects and concepts; and the spatial location of entities.	life by fishing from the end of a wharf"	
Emotion of reader	The inference is the emotion that the reader experiences when reading a text.	"...on removing the fork the eye came with it"	The reader is disgusted.
Author's intent	The inference is the author's attitude or motive in writing.	The entire passage	Bierce wants to lambaste workaholics.

Table 21: Knowledge-based inferences from "How Leisure Came". *Note*: From Constructing inferences during narrative text comprehension (Graesser et al. 1994: 375).

Further taxonomies of inferences in psycholinguistics and discourse processing are proposed in Rieger (1975), Clark (1977), Harris and Monaco (1978), Nicholas and Trabasso (1980), van Dijk and Kintsch (1983), Singer (1988), Magliano and Graesser (1991), Graesser and Kreuz (1993) as well as Kintsch (1993).

Above, we mentioned three carriers of implicit meaning which can be decoded by means of inferencing. These carriers are presuppositions, entailments, and implicatures.

In general, presupposition can be defined as information which the speaker of an utterance assumes to be already known by its addressee in order for an utterance to be semantically meaningful and appropriate in a given context (Fillmore 1975; Allwood, Andersson/Dahl 1977; Potts 2015).

Two notions of presuppositions – logical (also semantic and conventional) and pragmatic – have been established so far. Logical (semantic) presupposition, which has been already addressed in the Introduction, was first proposed by Frege (1892/2011) and is defined as a condition that must be true in order for a sentence in a context to have a truth value and be meaningful as exemplified by (3.23) in Section 3.3.1. Thus, the truth of presupposition is decisive for the truth of the sentence. If the presupposition fails, it leads to a truth-value gap of the sentence. That is, the latter is neither true nor false. Another definition of logical presupposition is provided in Keenan (1971: 45), stating that "a sentence *S* logically presupposes a sentence *S'* just in case *S* logically implies *S'* and the negation of *S*, *~S,* also logically implies *S'* ", attributing presupposition to sentences or propositions by this means.

One of the particularities of the logical presuppositions is that they "are part of the encoded meanings of specific words and constructions, called presupposition triggers" (Potts 2015:

169). Kiparsky and Kiparsky (1970), for example, mention factive verbs as presupposition trigger such as in (5.17).

(5.17) John *realized* that he had no money.

Presupposition: John had no money.

Factive verbs, however, are not the only possibility to signal presuppositions verbally. Some other lexemes and syntactic constructions have this ability as well. These are addressed, among others, in Potts (2015: 171), Levinson (1983), and Beaver and Geurts (2014) and summarized in Table 22.

Presupposition trigger	**Example**	**Study**
Aspectual predicates like *begin*, *continue, start, finish, leave, enter, cease,* and *stop*	China *has stopped* stockpiling metals. *Presupposition*: China used to stockpile metals.	Karttunen (1973), Simons (2001), Abusch (2002), Abrusán (2011).
Attitude predicates like *know*, *realize*, *to be sad that*, *to be aware*, and *regret*	Berlusconi knows that he is signing the end of Berlusconism. *Presupposition*: Berlusconi is signing the end of Berlusconism.	Kiparsky/Kiparsky (1970), Karttunen, (1973, 1974), Heim (1992), Abusch/Rooth (2004), Beaver (2001).
(In)definite determiners and demonstratives	The Prime Minister of Trinidad and Tobago stood up and wagged his finger. *Presupposition*: Trinidad and Tobago have a (unique) prime minister.	Strawson (1950, 1952/2011), Frege (1892/2011), Russell (1905/1988, 1957), Karttunen (1976), Kamp (1981), Heim (1982, 1983), Prince (1981), Roberts (2003), Elbourne (2005, 2008), Schwarz (2009), Schoubye (2009).
Pronouns like *his*, *her*, *our*, *their*	The farmer beats *his* donkey. *Presupposition*: The farmer has a donkey.	Karttunen (1976), Kamp (1981), Prince (1981), Heim (1982, 1983, 1990), Elbourne (2005).
Proper names	The author is Julius Seidensticker. *Presupposition*: Julius Seidensticker exists.	Prince (1981), van der Sandt (1992).
Quantifier Domains	I have written to every headmaster in Rochdale. *Presupposition*: There are headmasters in Rochdale.	Cooper (1983), Gawron (1996), Abusch/Rooth (2004), Roberts (2004).
Sortal Restrictions	Julius is bachelor. *Presupposition*: Julius is an adult male.	Thomason (1972).
Comparisons and contrasts marked by stress, additive particles like *too*, *also*, *in return*, *back*, and *either* and comparative constructions like *as much as*, *than*	Jimmy is as unpredictably gauche as Billy. *Presupposition*: Billy is unpredictably gauche.	Karttunen (1974), Heim (1992), Cohen (2009).

Adjunct clauses headed by prepositions like *before, during, whenever, as, since, while* and *after*	The dude released this video before he went on a killing spree. *Presupposition*: The dude went on a killing spree.	Heinämäki (1972, 1974), Beaver/Condoravdi (2003).
Appositives	The Proto-Harrappans, who flourished 2800-2650 B.C., were great temple builders. *Presupposition*: The Proto-Harrappans flourished 2800-2650 B.C.	Potts (2002 a,b), Schlenker (2010, 2009).
Clefts	It was Jesus who set me free. *Presupposition*: Somebody set me free.	Halvorsen (1980), Atlas/Levinson (1981), Soames (1982), Delin (1992, 1995),
Discourse particles like *even, just* and *only*	Only Alex does his homework. *Presupposition*: Alex does his homework	von Fintel (1999), Büring/Hartmann (2001), Beaver/Clark (2008).
Implicative verbs like *manage* and *fail*	John managed to open the door. *Presupposition*: John tried to open the door.	Karttunen (1971), Karttunen/Peters (1979).
Intonational contours, including topic and focus accents and verum focus	HE set me free. *Presupposition*: Somebody set me free.	Jackendoff (1987), Büring (1997), Geurts/van der Sandt (2004).
Evidentials	Presumably/probably it's the knave that stole the tarts. *Presupposition*: a) There is a (salient and identifiable) knave; b) There were (salient and identifiable) tarts; c) Somebody stole the tarts.	McCready (2005), Sauerland/Schenner (2007).
Manner adverbs like *quickly*	Jamie ducked quickly behind the wall. *Presupposition*: Jamie ducked behind the wall.	Abbott (2000).
Iteratives like *again*, *anymore*, *another time*, *to come back*, *restore*, *repeat*, and *for the nth time*	You can't get gobstoppers anymore. *Presupposition*: You once could get gobstoppers.	Levinson (1983).
Verbs of judging like *accuse* and *criticize* (this type of presupposition is a matter of controversy, as is not attributed to a speaker (see Wilson 1975)	Agatha accused Ian of plagiarism. *Presupposition*: (Agatha thinks) plagiarism is bad.	Fillmore 1975
Questions	Who is the professor of linguistics at MIT? *Presupposition*: Someone is the professor of linguistics at MIT	Katz (1972), Lyons (1977).
Counterfactual conditionals	If Hannibal had only had twelve more elephants, the Romance languages would not this day exist. *Presupposition*: Hannibal didn't have twelve more elephants.	Levinson (1983).

Table 22: Alleged presupposition triggers as listed in Potts (2015), Beaver and Geurts (2014), and Levinson (1983).

With respect to the CD task, presupposition triggers represent a valuable means for automatic processing. Consider the following example (5.18):

(5.18) a. Later in the day, Ukraine's security authorities reportedly said they *regained control of the two airports*.

b. There was *an attempt to seize the airports*, but we have localized those attempts. (ID 92)

The trigger *regained* in (5.18a) signals the existence of presupposition in the utterance, namely that Ukraine's security authorities had lost control of the two airports. However, (5.18b) utters that there was only an attempt to seize the airports, inferring that no control of the airports has ever been lost. Here we observe a lexical contradiction arising from uttering sentences of differing factuality.

Another particularity of logical presuppositions is their independence from the conventional meaning expressed by a sentence and as a consequence, their uncancelability under negation (5.19a), modals (5.19b), question (5.19c), and conditionals (5.19d).

(5.19) a. Ann stopped smoking.
Ann *didn't* stop smoking.
Presupposition: Ann smoked.

b. Ann stopped smoking.
Ann *must* stop smoking.
Presupposition: Ann smoked.

c. Ann stopped smoking.
Has Ann stopped smoking?
Presupposition: Ann smoked.

d. Ann stopped smoking.
If Ann stopped smoking, *then* …
Presupposition: Ann smoked.

While negation, for instance, affects the truth value of the sentence, modals change the modality of the sentence, and question modifies the speech act (an assertion has been changed to a question), the presuppositions remain stable and unchangeable. These tests or justifications are usually applied in order to distinguish between entailment (see below) and presupposition.

Another kind of presupposition to be addressed here is that of pragmatic presupposition. The notion of pragmatic presupposition has been developed by Stalnaker (1970, 1973, 1974, 1978) and is defined as "what is taken by the speaker to be the common ground of the participants in the conversation" (Stalnaker 1978: 321). A similar definition is provided in Givón (1979: 50), stating that pragmatic presupposition can be "defined in terms of assumptions the speaker makes about what the hearer is likely to accept without change". In both definitions, the speaker is represented as an "indicated source" of presuppositions (Brown/Yule 1991: 29). Potts (2015: 169), in turn, defines pragmatic presupposition as "the preconditions for linguistic interaction (for example, the mutual public knowledge that we are speaking the same language), the norms of turn-taking in dialogue, and more

particularized information about conversational plans and goals" (Potts 2015: 169). Pragmatic presupposition has been also addressed in Lakoff G. (1971), Lakoff R. (1971), Fillmore (1975), and Vennemann (1975). For current elaborations on pragmatic presupposition, see Stalnaker (1998) and Simons (2003).

According to Potts (2015: 169) "the clearest instances of pragmatic presuppositions are those that cannot easily be traced to specific words or phrases, but rather seem to arise from more general properties of the context and the expectations of the discourse participants". The pragmatic presupposition is not a part of the propositional content of a sentence, but rather something which is taken for granted by the speaker. That is, the speaker assumes that this information belongs to the common ground of both him and his addressee. In the case that the addressee does not possess the required knowledge, i.e. is not able to appropriately recognize/establish/accept the presupposition, the addressee can signal this to the speaker. By this means, a new information is communicated. Moreover, the addressee can accept the presupposition as the default (or given) and add it to the common ground. This is referred to as accommodation (Lewis 1979). Approaches to the modeling of common ground are addressed in Gauker (1998), Gunlogson (2001), and Farkas and Bruce (2010).

Dahlgren (1974), in turn, sees the difference between two kinds of presupposition in that that logical presupposition is a relation between two sentences, while pragmatic presupposition is the relation between a sentence and the beliefs of the speaker in order that a sentence may be used appropriately. Dahlgren (1974: 1) formulates this as follows: "the logical notion which takes presupposition to be a (logical) relation between two sentences, and the pragmatic notion which incorporates the beliefs of the speaker and aspects of the context of the speech act into the set of things presupposed by a sentence". It should, however, be noted that although the term logical (semantic) presupposition is used to clearly distinguish it from pragmatic presupposition, it is also pragmatic to some degree as logical (semantic) presuppositions must be evaluated on the common ground of the discourse participants (Potts 2015).

The notion of logical presupposition is often confused with the notion of logical entailment. Consider the following example (5.20):

(5.20) a. That person is a bachelor. (=*A*) (Kempson 1975: 48)

b. That person is a man. (=*B*)

It is evident, that (5.20a) implies (5.20b). A person who is a bachelor befits the definition of the term bachelor as a male. We observe a logical entailment here, as *B* logically follows from *A* (or *B* is entailed by *A*, or *A* entails *B*). In general, the relation of logical entailment between sentences is defined in terms of their truth values. Thus, for logical entailment, it

applies that it is necessary for *B* to be true if *A* is true (cf. Löbner 2013: 175). If *A* entails *B*, then *A* is referred to as the premise of the entailment, while *B* is its conclusion. Further, for logical entailment, it applies that it is impossible that *B* is false when *A* is true, and *A* cannot be true if *B* is false. Together with equivalence, contradiction, and contrariety, logical entailment represents the logical relation between sentences (Löbner 2013).

The main difference of entailment as compared to presupposition consists in the observation that it does not survive under the embeddings of the operators (in first turn negation). Thus, in the case of negating (5.20a), (5.20a) does not hold the implication (5.20b) and the entailment relation does not persist. Further, with respect to pragmatic presupposition, it can be observed that while pragmatic presupposition is a relation between a speaker and a sentence, entailment represents a (logical) relation between sentences.

> "A [pragmatic - NKB] presupposition is something the speaker assumes to be the case prior to making an utterance. Speakers, not sentences, have presuppositions. An entailment is something that logically follows from what is asserted in the utterance. Sentences, not speakers, have entailments" (Yule 1996: 25).

With respect to the automatic CD, the recognition of entailment relation between the sentences can positively contribute to finding related sentences which are potentially contradictory as, e.g., in the case of (5.21). For this, Harabagiu et al. (2006) and de Marneffe et al. (2008) suggest first removing the explicit contradiction cues and then checking the sentences (*T* and *H*) for entailment (Section 2.1).

(5.21) a. Ross Aimer, a former pilot with United Airlines, told Al Jazeera it was highly *unusual* that air traffic control *would lose contact with an aircraft without communication* from the crew.

b. In that area of the world, over Vietnam, there is *sporadic radar coverage*, to begin with," he said. (ID 005)

Sporadic radar coverage in the area over Vietnam (5.21b) entails that a contact between an aircraft and the air-traffic controller cannot always be properly established. This contradicts the information provided in (5.21a).

The last kind of implication that we want to address here is the concept of implicature. The notion of implicature will only briefly be described here as due to its complexity, it won't be considered for the development of the CD system.

Implicature was introduced to the philosophy of language by Grice in his article, *The Causal Theory of Perception* (1988), and further treated thoroughly in *Logic and Conversation* published in *Studies in the Way of Words* (1967/1989). In these studies, the scientist expounds his observation that the actual meaning of an utterance is often not directly communicated but inferred by the text recipient/hearer, depending on his particular background knowledge and on the context in which the utterance is expressed. Based on this observation, Grice

suggested distinguishing between “what is said” on the one hand and “what is meant” on the other hand – an aspect which gave rise to distinguishing between pragmatics and semantics.

Grice (1967/1989) distinguishes between two kinds of implicature – conventional and conversational. The conventional implicature can be determined based on the conventional lexical meanings of the words; “[…] in uttering a sentence S, a speaker implies that p is the case if, by having been uttered, S suggests as its conclusion p, without p having been literally said. If the conclusion rests exclusively on the conventional meaning of the words and grammatical constructions that occur in S, then the conclusion is called a ‘conventional implicature” (Bussmann 2006: 221). Bussmann (2006) further points out that since Karttunen and Peters (1979), most presuppositions are treated as conventional implicatures.

In contrast to conventional implicature, conversational implicature refers to a message conveyed by a sentence which is beyond its literal meaning. That is, conversational implicature is not a part of the conventional lexical meanings of the words either; it is not determined by the structure of the sentence. In sentences with different structure but which are semantically related, implicature remains (*I don't like apples - I hate apples – I won’t eat them*). Conversational implicatures, in turn, are not explicitly expressed or encoded but are communicated above the literal meaning. They have to be decoded or “understood” by the listener/reader.

The main mechanism which drives or initiates the inference of conversational implicature in communication, according to Grice (1967/1989), is the Cooperative Principle. Grice (1967/1989: 26) states: “Make your conversational contribution such as is required at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged”. The Cooperative Principle is regulated by a number of interaction rules, the so-called conversational maxims. These include according to Grice (1967/1989: 26-27):

- The Maxim of Quality:
    a. Do not say what you believe to be false;
    b. Do not say that for which you lack adequate evidence;
- The Maxim of Quantity:
    a. Make your contribution as informative as is required (for the current purposes of the exchange);
    b. Do not make your contribution more informative than is required;
- The Maxim of Relation:
    a. Be relevant;
- The Maxim of the Manner:
    a. Avoid obscurity of expression;

b. Avoid ambiguity;
c. Be brief (avoid unnecessary prolixity);
d. Be orderly.

Consider the following situation:

(5.23) Participant 1: I am out of petrol. (Grice 1967, 1989: 32):
Participant 2: There is a garage round the corner.

Assuming that Participant 2 is cooperative (the Cooperative Principle), Participant 1 can infer that the utterance of Participant 1 is relevant to her/his problem that she/he is out of petrol. As both participants share the same knowledge about the garage, Participant 1 understands the utterance of Participant 2 so that the latter wants to provide information that there is a garage nearby which is open and where Participant 1 can purchase petrol.

In contrast to presupposition, conversational implicature can be canceled by the context. Moreover, unlike presuppositions and entailments, conversational implicatures are not truth-functional. Their "truth" depends on the context as well as on the ability of the conversation participants to cooperate. That is, conversational implicatures are a matter of interpretation. They are said to be based on stereotyped expectations and, therefore, can be often understood or interpreted incorrectly.

5.3 Giessen Annotated Corpus of Contradictions in News Texts

5.3.1 Survey 3: Typology Validation, Results, and Evaluation

To validate the nine types of contradictions identified, including explicit-negation, explicit-opposition, explicit-numerical, explicit-lexical, implicit-negation, implicit-opposition, implicit-numerical, implicit-lexical, and implicit-factual, another survey – Survey 3 – has been conducted. The same three participants of Survey 2 took part in Survey 3.

The participants were instructed about the purpose of the study and were given descriptions of the contradiction types. Their task was to decide on the type (or multiple types) of each of the 311 contradictions collected in Survey 1 and Survey 2. For processing of the questionnaires, the participants were allotted a time period of one month.

The inter-annotator agreement has been computed to evaluate the degree to which the raters made the same decision on a certain type/or types of contradiction based on the descriptions provided. For this purpose, Fleiss's Kappa and Krippendorff's Alpha measures were again preferred. For computing the scores, the online tool developed by J. Geertzen has been applied (see Section 5.1.2). The results of the survey before and after the discussion of the variables relatedness, contradiction cue, and contradiction type are presented in Table 23.

Inter-rater agreement score	Fleiss			Krippendorff		
	A_obs	A_exp	Kappa	D_obs	D_exp	Alpha
Before discussion						
Relatedness	0.959	0.525	0.914	0.04	0.475	0.916
Contradiction cue	0.93	0.348	0.893	0.068	0.652	0.896
Contradiction type	0.9	0.183	0.878	0.097	0.818	0.881
After discussion						
Relatedness	1.0	0.523	1.0	0	0.478	1.0
Contradiction cue	1.0	0.348	1.0	0	0.653	1.0
Contradiction type	1.0	0.182	1.0	0.0	0.819	1.0

Table 23: The scores of inter-rater agreements on the type of relatedness, contradiction cue, and contradiction type, before and after discussion, computed with Fleiss's Kappa and Krippendorff's Alpha.

As can be seen in Table 23, although the task of type assignment is a complex task, the computed scores – 88% for both Fleiss's Kappa and Krippendorff's Alpha – indicated a high agreement among the raters on the contradiction types. Hereby, the raters showed a slightly better agreement on the kind of relatedness between the parts of contradictions (computed at 91% with Fleiss's Kappa and 92% with Krippendorff's Alpha) than on the kind of contradiction cue (computed at 89% with Fleiss's Kappa and 90% with Krippendorff's Alpha). In order to achieve agreement on the assigned types in some way, an additional discussion with the raters took place. As the result of this discussion, a consensus between the raters on contradiction types could be achieved for all pairs of contradictions (Fleiss's Kappa and Krippendorff's Alpha score amounted to 1.0 (100%), see Table 23).

5.3.2 Corpus Annotation

The final step was the annotation of the collected contradictions along with their types, as well as news texts in which the contradictions occurred, including the metadata of the texts.

Annotation can be defined as a task of assigning original data with additional information. The corpus annotation of the types of contradictions serves two purposes. First, it aims at proving whether the typology was correctly constructed and the task for type identification/annotation is well defined, that is, whether the description of the types is adequate and clear and worth processing by computation. This step has been already performed as described in Section 5.1.2.1 by means of questionnaires. Second, annotation serves the purpose of providing the collected data in an appropriate standard unified format for use in other projects on contradictions in news texts as well as in general studies on contradiction. The aspect of sustainability is provided by these means.

For the annotation, the XML markup language has been chosen. The layout for the annotation is illustrated in Figure 4.

```
<?xml version="1.0" encoding="UTF-8"?>
<corpus>
        <topic id="n" title="">
                <articles>
                        <article article_id="n">
                                <meta>
                                        <title></title>
                                        <date></date>
                                        <url></url>
                                        <source></source>
                                 </meta>
                                 <text>
                                        <sentence n="1"></sentence>
                                        <sentence n="2"></sentence>
                                        <sentence n="n"></sentence>
                                </text>
                        </article>
                </articles>
                <contradictions>
                        <pair id="n" relatedness="" contradiction_cue="">
                                <sentence n="n" article_id="n"></sentence>
                                <sentence n="n" article_id="n"></sentence>
                        </pair>
                </contradictions>
        </topic>
</corpus>
```

Figure 4: XML annotation layout of the corpus.

All articles in the corpus <corpus> are grouped according to their topic <topic>. Each <topic>-tag has an id attribute, which begins with 01 and is subsequent for all articles in the corpus, and a title attribute, which is a subject dealt with in the articles. In total there are nine topics in the corpus. Each <topic>-tag further includes correspondent articles labelled with <articles> and a separate list of contradictions labelled as <contradictions>. Enumeration of the articles begins with 001 and is subsequent for all articles independent from the topic. Each article is provided with meta information (<meta>) such as title (<title>), date (<date>), url (<url>) and original source of the text (<source>) and a text (<text>) consisting of sentences (<sentence>). In turn, <contradictions>-tag includes pairs of contradictory

sentences. Each pair of contradictory sentences (<pair>) includes an id attribute, the type of contradiction according to the relatedness of its parts (<relatedness>) and contradiction cue (<contradiction_cue>). The enumeration of the contradiction IDs in the corpus begins at 000. Annotation of news texts and their information, as well as pairs of contradictory sentences, including metadata but excluding their type, was performed automatically. A Python script was written for this purpose. The types of contradictions were annotated manually. The Giessen Annotated Corpus of Contradictions in News Texts is located in Attachment C.

In total, the compiled corpus contains 311 contradictions. From these, 302 consist of two parts (sentences), while 9 of contradictions are represented by a single sentence. Further, 12 contradictions contain contradictory elements of two types and 299 contradictions contain contradictory elements of one type. These are the contradictions with the IDs 063, 179, 185, 222, 233, 235, 242, 244, 263, 267, 305, and 310. Further, 127 explicit and 196 implicit contradictions in the corpus are presented, which amount to 39% and 61% of the corpus, respectively. The distribution of contradictions, according to the kind of contradiction cues, is illustrated in Figure 5.

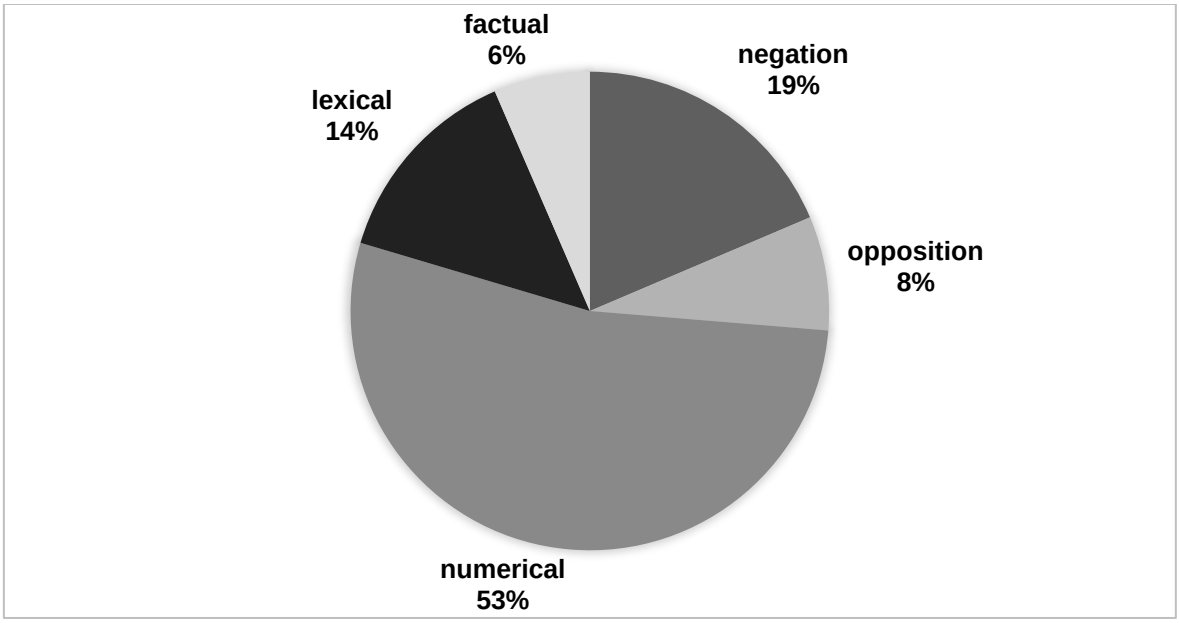

Figure 5: Distribution of contradictions in the corpus according to contradiction cues.

As can be seen in Figure 5, the contradictions arising from numerical divergencies (172 instances, or 53.3% of the whole corpus) and not negation as expected (60 instances, or 18.6% of the corpus) represent the most frequent kind of contradictions in the corpus. A possible explanation for this finding is that numerical contradictions – both explicit and implicit – could be more easily recognized by Survey 1 participants than the (explicit and implicit) negation contradictions. In turn, factual (21 instances, or 6.5% of the corpus) and opposition contradictions (25 instances which are 7.7% of the corpus) represent the less

frequent kind of contradictions in the corpus. In case of factual contradictions, this can be explained by the difficulty of their detection since world knowledge is required. Finally, contradictions arising from lexical incompatibilities are represented in the corpus with 45 instances (13.9%).

Figure 6 illustrates the distribution of contradictions according to their type. Contradictions of the explicit-numerical type (80 instances, or 24.8% of the corpus) and the implicit-numerical type (92 instances, or 28.5% of the corpus) most frequently occur in the corpus, again justifying the assumption that numerical divergencies, either explicit or implicit, are easily detected by a human. Such a frequent occurrence of numerical contradictions highly motivates the integration a component for dealing with numerical values in a system.

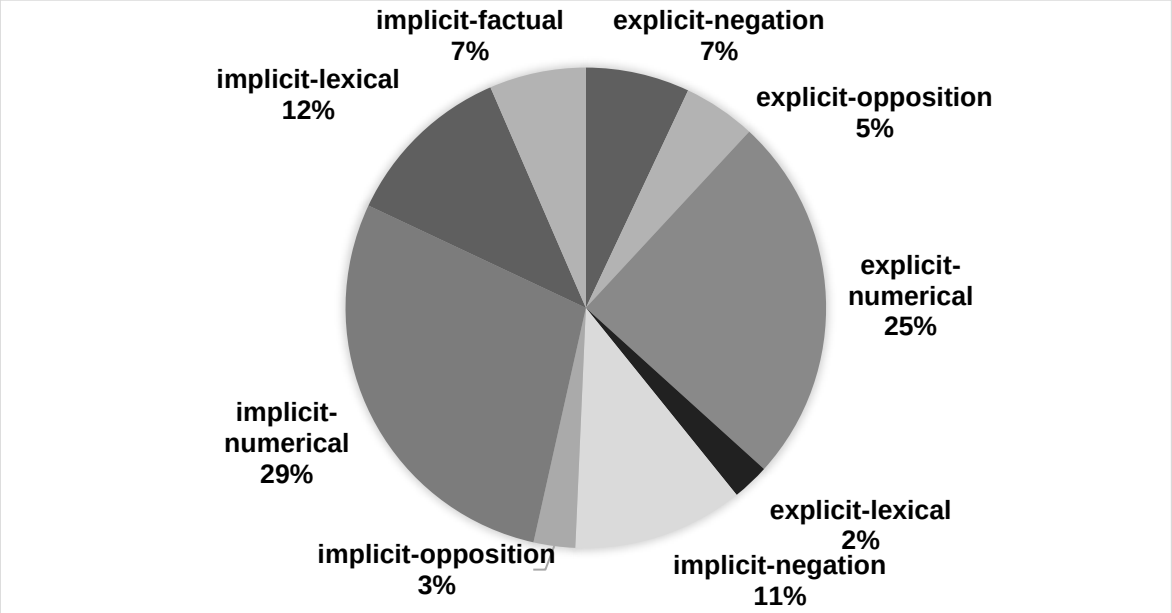

Figure 6: Distribution of types of contradictions in the Giessen Annotated Corpus of Contradictions in News Texts.

In contrast, contradictions of types explicit-lexical (8 instances or 2.5%) and implicit-opposition (9 instances or 2.8%) occur less frequently, which can be due to diverse reasons. Slightly more frequent in the corpus are contradictions of types explicit-opposition (16 instances or 4.9%), implicit-factual (21 instances or 6.5%) and explicit-negation (23 instances or 7%). Finally, contradictions of types implicit-negation and implicit-lexical are represented in the corpus in equal proportions, namely by 37 instances each, which constitutes 11.5% of the whole corpus.

5.4 Summary

The present chapter described the process of a corpus-based identification of realization mechanisms and the cues of contradictions occurring in news texts. By that, the focus was

set on features that can be processed by a machine for identifying a contradictory relation between two sentences.

For the purpose of typology construction, the dimension *relatedness of the contradiction parts* on one side and computationally oriented dimension *contradiction cue* on the other side, have been selected.

Based on the evidence in the collected newswire data, five kinds of contradiction cues have been identified. These include negation, opposition relation, numerical, factual, and lexical divergencies. Beside contradiction cues, implicit and explicit kinds of contradictions could be identified and their realization mechanisms described. That is, while explicit contradictions are characterized by a verbal similarity of their parts, implicit contradictions can be identified by revealing of implicit meaning carried by presupposition, entailment relation, and implicature first. Combining the two typology dimensions, nine types of contradictions occurring in newswire texts could be identified, including explicit-negation, explicit-opposition, explicit-numerical, explicit-lexical, implicit-negation, implicit-opposition, implicit-numerical, implicit-lexical, and implicit-factual. By that, explicit-numerical and implicit-numerical contradictions are the most frequent types, while explicit-lexical and implicit-opposition contradictions are less represented in the corpus. It should be noted that the constructed typology does not have a claim to be completed.

6 Conceptual Design of a CD System and Supporting Tools

The recognition of contradictions can be described in a simplified form as a process of finding of at least two declarative sentences or parts of a declarative sentence from one or more texts which have the same semantic content (proposition) and refer in the same respect and time to the same situation in the world as well as include elements that are contradictory or in contrary relation with each other. Though the task seems to be easy, it poses a challenge to an automatic CD, requiring a variety of natural language processing steps at lexical to discourse levels. The main aim of the present chapter is first to summarize the theoretical elaborations on contradiction from the previous chapters and then incorporate them into a conceptual design (what a system should do) of a CD system (Section 6.1). The conceptual design will serve as a basis for a physical design, which is how the system should be built (Chapter 7). Logical design, which refers to a buildup of a system with respect to a user, will stay beyond the scope of the study. Second, the aim of the present chapter is to introduce the reader to the state of the art of supporting tools for the CD task with a focus on English. The chapter begins with a discussion of NLP tasks at lexical, morphological, and syntax levels (Section 6.2.1). In the focus of the section are the tasks of tokenization and sentence splitting (Section 6.2.1.1), stop word detection and removing (Section 6.2.1.2), part-of-speech tagging (Section 6.2.1.3), stemming and lemmatization (Section 6.2.1.4) as well as parsing and chunking (Section 6.2.1.5). The existing techniques for meaning processing at semantic, pragmatic, and discourse levels are addressed in Section 6.2.2, including semantic role labeling (Section 6.2.2.1) and textual entailment recognition (Section 6.2.2.2). How and to what degree the sentence meaning construction and interpretation can profit from the properties of a text is discussed in Section 6.2.2.3. Here, the notions of cohesion and coherence (Section 6.2.2.3.1) and state-of-the-art tools for anaphora (or coreference) resolution (Section 6.2.2.3.2) are of particular interest. The NLP tasks relevant to a CD task include the negation and modality processing (Section 6.2.3.1), sentiment analysis (Section 6.2.3.2), named entity recognition (Section 6.2.3.3) as well as temporal (or time) processing (Section 6.2.3.4), and measuring semantic textual similarity (Section 6.2.3.5). Section 6.2.4, in turn, outlines the existing approaches to meaning representation, which range from a logical form to the view of meaning as a word embedding. Finally, Section 6.2.5 introduces the prominent computational sources of knowledge, including lexical resources (Section 6.2.5.1) and ontologies (Section 6.2.5.2).

6.1 Conceptual Design

Considering the theoretical and empirical elaborations on contradictions in the previous chapters, the following tasks for detection of contradictions occurring in single or multiple news texts are prior for a system. Thus, the system should be able

- To process newswire textual data considering its particular characteristics including frequent use of direct and indirect speech, particular language use characteristics such as the frequent occurrence of noun phrases, numbers, and figures. The system should also be able to prepare the unstructured data for further applying the NLP tasks. The latter is required to gather information (i.e. lexical meaning relations, thematic roles), supporting the task of contradiction detection;

- To identify temporal reference of the described event to justify the conditions of the contradiction according to the ontological view on contradiction as addressed by Aristotle;
- To determine the truth value of the sentences for justifying the semantic view of contradiction as addressed by Aristotle;
- To process not only quantification but also the time, date, quantity, and quality expressions, i.e. to identify equivalent expressions, such as two hours to be 120 minutes;
- To cope with modality, that is to classify the described event with respect to the alethic (“necessary”, “possible”, “impossible”), deontic (“obligatory”, “permitted”), and the epistemic (“know”, “believe”, “imagine”) kinds of modality. Further, the system should be able to recognize the event’s factuality status (happened, not happened, underspecified);
- In the context of the previous point, the system should be able to process negation, including identifying the scope of negation, and double negation. The recognition of negated events is also required for detection of the contradictions arising from negation;
- To find parts of a contradiction/contrariety, which can be single sentences, paragraphs, or whole texts.
- To process the text and identify the information implicitly communicated in the text. This requires the ability of logical (deduction and induction) and pragmatic inferencing. The latter includes the recognition of presupposition (logical and pragmatic), textual entailment, and implicatures (in first turn conventional). In the context of inferencing, the ability of a system to address and process different kinds of world knowledge is required;
- In order to identify fake contradictions, the system should be able to conduct the disambiguation of word meanings and recognize events which are verbally equivalent or similar but are not co-referent, that is do not refer to the same object in the real world;
- To recognize vague terms in order to identify borderline contradictions.

6.2 Supporting Tools and Methods

6.2.1 Processing at Lexical, Morphological and Syntax Levels

6.2.1.1 Tokenization and Sentence Splitting

From a computational point of view, a raw text is a single sequence of symbols or characters, termed a string. Thus, before the textual data can be further analyzed by a machine

applying diverse tools, it has to be broken up into some significant units, or constituents, which can be words and punctuation (tokens[22]) or sentences. The process of splitting the text into sentences is referred to as sentence splitting and splitting into the words as tokenization.

Despite the simplicity of the task at first sight, there is a number of approaches to tokenization and sentence splitting which follow different principles. No standard approach exists at the present time. The preferred technique for tokenization follows the whitespaces between words. This approach, however, can be unreliable in the case of the compound words such as *natural language processing* (cf. Jackson/Moulinier 2002: 10; Krüger-Thielmann/Paijmans 2004: 357). Also, in case of abbreviations and acronyms, it is often not clear whether a punctuation mark is a part of a word such as in *Dr.* or not.

In turn, the preferred approach to sentence splitting in English follows the punctuation marks such as ".", "?", "!", "…", which conventionally signal the end of a sentence. A problematic issue for this approach is posed by the cases where the punctuation marks are the part of abbreviations and acronyms such as in *Dr.* and *M.B.A.* For reliable output, such cases have to be taken into consideration. Besides the orientation on whitespaces, another approach searches for capitalized words which usually mark the beginning of a sentence. This approach faces difficulties because some words, e.g. names, are always capitalized. A possible solution would be to work with predefined lists of rules as exceptions (cf. Jackson/Moulinier 2002: 9-10).

The tasks of tokenization and sentence splitting are usually built-in in most of the state of the art NLP tools so that the use of external tokenization tools is not required. Some algorithms, however, such as Charniak's parser (Charniak 2000) require external sentence splitting. The most robust state-of-the-art sentence splitter has been proposed by Reynar and Ratnaparkhi (Reynar/Ratnaparkhi 1997). The sentence splitter could achieve an accuracy of almost 99% evaluated on the Wall Street Journal dataset.

---

[22] The terms token and type (German: Vorkommnis and Typ) have been first introduced by Peirce (1906: 505-506) and are defined as follows: "A common mode of estimating the amount of matter in a MS, or printed book is to count the number of words. There will ordinarily be about twenty thes on a page, and of course they count as twenty words. In another sense of the word 'word,' however, there is but one word 'the' in the English language; and it is impossible that this word should lie visibly on a page or be heard in any voice, for the reason that it is not a Single thing or Single event. It does not exist; it only determines things that do exist. Such a definitely significant Form, I propose to term a Type. A Single event which happens once and whose identity is limited to that one happening or a Single object or thing which is in some single place at any one instant of time, such event or thing being significant only as occurring just when and where it does, such as this or that word on a single line of a single page of a single copy of a book, I will venture to call a Token." Besides letters and words, tokens and types can be also applied to sentences, paragraphs, etc.

A popular at present NLTK (Bird, Klein/Loper 2015), which is a collection of libraries, programs, and resources for NLP using the programming language Python, provides several functions for tokenization. These include (1) *split()*, which follows the whitespaces between words and splits, (2) a Treebank tokenizer *word_tokenize()*, which splits a string into tokens, including splitting most punctuation from adjoining words as well as verb contractions (e.g. *I'm*, *won't*) and the Anglo-Saxon genitive of nouns (e.g. children*'s*) into their component morphemes (e.g. *I'm* to *I* and *'m*), (3) *wordpunct_tokenize()*, which functions similarly as *word_tokenize()* but additionally splits punctuation marks from the adjoining word, as well as (4) a *RegexpTokenizer()* tokenizer which is based on regular expressions. Sentence splitting, in turn, can be performed by means of the function *sent_tokenizer.tokenize()*.

Additionally, the Stanford CoreNLP package, which provides a number of essential tools for NLP developed by the Stanford Natural Language Processing Group, contains a TokenizerAnnotator component for splitting a text into tokens and a WordsToSentencesAnnotator component for sentence splitting.

6.2.1.2 Stop Word Detection and Removing

Stop words can be defined as items with little or no semantic content or items which are irrelevant for the semantics of a certain document. Usually, one counts prepositions such as *on*, *at*, *in* as well as definite and indefinite articles *a* and *the* to the stop words. Also, words of other word classes such as adjectives can be regarded as stop words if they are not important for representing the content of the document (Gödert, Lepsky, and Nagelschmidt 2012; Schmolz 2015).

For the task of stop words detection, pre-defined word lists are required. These can be generated manually or automatically. The NLTK package e.g. provides a corpus of stop words which can be accessed by importing the module *corpus*. The NLTK stopwords corpus includes high-frequency words with little or no semantic content such as *the*, *to*, and *also*.

It should be considered that, depending on the purpose of the project, by detection and removing of stop words, one can achieve either positive or negative effects. From the positive point of view, besides the increasing reliability of the results, removing the stop words can reduce the required memory capacity to 30-50% (Schmolz 2015). Further, Schmolz (2015) points out that the detection and elimination of stop words can improve the precision and recall scores. A negative effect of the elimination of stop words can be observed in case of processing quotations such as *To be or not to be*, which consequently, will not be found. Abbreviations such as *THE* (Times Higher Education) will be lost as well. Finally, removing stop words can have negative consequences for the performance of the anaphora resolution systems (see Section 6.2.2.3 for more information on anaphora resolution).

6.2.1.3 Part-of-speech Tagging

Part-of-speech tagging (henceforth POS tagging) is a task of an automatic identification and assignment of a part-of-speech label (or POS tag) such as noun, verb, adjective, adverb, etc., to a token – a word or another kind of token (e.g. punctuation mark, number). This task is of importance, e.g., for further NLP tasks such as lemmatization and recognition of named entities (Sections 6.2.1.4 and 6.2.3.3, respectively).

POS tagging is not trivial, as many words are ambiguous, i.e. can be assigned with different pos tags. Thus, e.g., depending on the context, the word access can be a noun as in *provide an access* or a verb as in *to access the Internet*.

There is a number of approaches to POS tagging including rule-based, probabilistic, and transformation-based (Uryupina/Zanoli 2016). While rule-based taggers assign pos tags following a set of manually written rules, probabilistic taggers determine the correct pos tag based on a training corpus manually annotated with pos labels in order to get the most probable tag. Transformation-based approaches, in turn, combine rule-based and probabilistic approaches. The accuracy of the state of the art taggers achieves 97%. One should, however, note that such a high accuracy can be achieved by applying the taggers to standardized textual data (such as news texts). In case of the POS tagging of computer-mediated data (e.g., forum, blog) more sophisticated approaches are required.

To the widely used pos taggers for English at present belong TreeTagger, Stanford Tagger as well as TnT Tagger.

The TreeTagger (Schmid 1994, 1995), developed at the University of Stuttgart, is considered to be one of the most widely used taggers. Besides POS tagging, TreeTagger can be applied to the task of lemmatization. TreeTagger is a probabilistic pos-tagger that uses decision trees. It can be compared with the n-gram tagger described in Church (1988) and Kempe (1993). Either TreeTagger or n-gram tagger model the probability of a tagged sequence of words. TreeTagger, however, estimates the transition probabilities using a binary decision tree, whereas the n-gram tagger follows a formula based on the maximum likelihood estimation (Schmid 1994). TreeTagger for English uses 45-tag Penn Treebank tagset (Marcus et al. 1993). One of its main advantages is that it has been developed for a number of languages including, among others, English, German, French as well as Chinese, Swahili, Latin, and Greek. TreeTagger can be used as an off-the-shelf tool or can be accessed within the NLTK Python package but also by other programming languages such as R, Perl, Java, Ruby, and within the UIMA architecture. The guidelines for the integration of TreeTagger into NLTK can be found on the official website of TreeTagger. The tool has been evaluated on the Penn Treebank dataset and achieved a 96.36% accuracy. It is an open

source for academic use and can freely be downloaded from the official website of the project.[23]

TnT, or Trigrams'n'Tags (Brants 2000), is a statistical part-of-speech tagger based on the implementation of the Viterbi algorithm for second-order Markov models and has been developed at Saarland University. One of the advantages of the tool is that it additionally provides the possibility for training one's own tagging model on a variety of languages and tag sets. Moreover, it provides different options for treating unknown words. The tool can be used as an off-the-shelf tool or within the NLTK package. TnT has been evaluated on Penn Treebank Wall Street Journal, where it could achieve 96.46% accuracy on all tokens and 85.86% on unknown words. Besides the Penn Treebank (English), the tagger has been evaluated with 94.5% accuracy on the Susanne Corpus (English) and with 96.7% accuracy on the NEGRA corpus (German). TnT is an open source program for academic use (non-commercial research use) and can be downloaded for application on German and English languages after signing the license agreement.[24]

The Stanford Tagger (Toutanova/Manning 2000; Toutanova et al. 2003) is another pos tagger which is available as a PosTaggerAnnotator component of the Stanford CoreNLP package. The system requires the prior installation of Java 1.8+. Additionally, in order to use the already trained tagger, 60 MB to 200 MB of memory is required. The system also provides a possibility to train one's own tagger. For this purpose, depending on how complex a model will be, at least 1 GB of memory is required. The tagger can be trained on any language. One of the main advantages of the system for those who are not familiar with Java is that the Stanford Tagger can be applied within the Python-based NLTK package. The Stanford Tagger has been evaluated on the Penn Treebank Wall Street Journal and achieved 97.24% accuracy on all tokens and 89.04% on unknown words. The application was further improved by Manning (2011). The improved version of the Stanford Tagger, evaluated on the Penn Treebank WSJ, could achieve 97.32% accuracy on all tokens and 90.79% accuracy on unknown words. It is an open source and can be directly downloaded from the website of the Stanford Natural Language Processing Group.[25]

6.2.1.4 Stemming and Lemmatization

In textual data, a word usually occurs in different grammatical forms such as e.g. *thinks*, *thought*, and *thinking*. Depending on the purpose of the project, it can be reasonable to automatically reduce the flectional or derivational form of the words to their base form

[23] http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/
[24] http://www.coli.uni-saarland.de/~thorsten/tnt/
[25] https://nlp.stanford.edu/software/tagger.shtml

(*think*s, *thought*, and *thinking* → *to think*). For this reason, the tasks of stemming or lemmatization can be applied, which treat the flections in two different ways. While lemmatization uses vocabulary and conducts morphological analysis of the words to return a vocabulary form of the word, the so-called lemma (e.g. *thinking – to think*), stemming simply strips the ends of the words (e.g. *think~~ing~~*), by this means often achieving the correct result.

According to the underlying approach, the stemming algorithms can be classified in rule-based (e.g. the Lovins stemmer (Lovins 1968), Porter's stemmer (Porter 1980), the Paice/Husk Lancaster stemmer (Paice 1990) as well as the Dawson stemmer as extension of the Lovins stemmer), statistical (e.g., n-gram stemmer and a stemmer based on a hidden Markov model (Melucci/Orio 2003)) and hybrid stemmers (e.g. the Krovetz stemmer (Krovetz 1993) and a corpus-based stemmer (Xu/Croft 1998)). For an overview of existing stemmers, see Jivani (2011).

The most common rule-based algorithms for stemming are Porter's Stemmer (Porter 1980) and the Lancaster stemmer, also known as the Paice/Husk stemmer (Paice 1990). Porter's stemmer is one of the first stemmers and is also the most preferred one, which can be explained by "its simplicity and elegance" and as it "yields results comparable to those of the more sophisticated algorithm" (Baeza-Yates/Ribeiro-Neto 2011: 227). Porter's stemmer is a rule-based algorithm and processes stemming in five phases of word reductions, which are applied sequentially. Each phase includes different conventions for the selection of a rule from a set of rules. Compared to other stemmers, Porter's stemmer achieves the best stemming results.

The Lancaster stemmer, in turn, is an iterative stemmer which is known as strong and aggressive. It is based on a single table of 120 rules indexed by the last letter of a suffix. Each rule in the table specifies either the deletion or the replacement of an ending. On each iteration, the algorithm tries to find the appropriate rule according to the last letter of the word. If no rule is found, the algorithm terminates. The disadvantage of the Lancaster stemmer is that the output is usually overstemmed so that due to the extreme shortness of the words, the latter is difficult to comprehend. The use of the Lancaster stemmer, however, can be of advantage when the processing of a large amount of textual data is involved. Both stemmers can be directly used within NLTK. For lemmatization, the NLTK WordNet lemmatizer provides a simple possibility to perform lemmatization (for more information on WordNet, see Section 6.2.5.1). The outputs of Porter's and the Lancaster stemmers, as well as a WordNet lemmatizer, applied on the quote from Lewis Carroll's *Alice in Wonderland*, are as follows:

- Sample text:
  Why, sometimes I've believed as many as six impossible things before breakfast.

- Porter's stemmer:
  Whi , sometim I ' ve believ as mani as six imposs thing befor breakfast .
- Lancaster stemmer:
  why , sometim i ' ve believ as many as six imposs thing bef breakfast .
- WordNet lemmatizer:
  Why , sometimes I 've believe a many as six impossible thing before breakfast.

Also, TreeTagger (Section 6.2.1.3) can be applied to generate lemmas for words in the text data. The Stanford CoreNLP package's component for lemmatization is the MorphaAnnotator.

6.2.1.5 Parsing and Chunking

Parsing in NLP refers to the automatic syntactic analysis of a sentence in accordance with a particular grammar theory. The term has to be distinguished from the term parsing in computer science, which is defined as a process of dividing the source code into meaningful parts for processing by a machine. In literature, often the term shallow, or partial, parsing is used. Shallow parsing, or chunking, is the process of partial identification of sentence constituents or chunks (usually noun phrases for the task of NER), based on the pos labels of the words, without following complex grammar rules and without determining the hierarchical relations between the constituents. Therefore, the difference between parsing and shallow parsing lies in the depth of the sentence analysis.

Traditionally, parsing in NLP follows two grammar theories for representing the syntactic relations – the constituency and the dependency grammars. The idea behind the constituency grammar is that the syntactic structure of a sentence can be represented hierarchically as a sequence of words or groups of words that function as a single common unit (e.g., noun phrase, verb phrase). Dependency grammar, in turn, assumes that there exist direct dependencies between words, e.g., the subject and its object depend on the main verb. However, contextual ambiguity poses a problematic issue in the parsing.

The traditional approach for modeling the constituent structure of English and other languages is the context-free grammar (CFG), or phrase-structure grammar. A CFG consists of a set of rules (or productions) expressing the ways in that symbols (e.g., V for verb, NP for noun phrase) can be ordered using a lexicon of words and symbols (cf. Jurafsky/Martin 2008: 387). The output of a syntactical analysis is usually presented as a parse tree or, more compactly, as a bracketed notation.

Different approaches for parsing by CFG have been proposed so far, including bottom-up, top-down, a combination of top-down parsing with bottom-up filtering, as well as chart

parsing. For more information on these approaches, see Jurafsky and Martin (2008), Carstensen et al. (2010), Lobin (2010), and Bird et al. (2015).

Among others, the preferred constituency and dependency parsers publicly available are the Collins parser (Collins 2003), Charniak's parser (Charniak 2000) as well as the Berkley parser (Petrov 2006). These parsers have been evaluated on the Penn Treebank Wall Street Journal dataset and achieved 87-91% as a computed F-measure. Further available parsers are the MaltParser,[26] the ParserAnnotator, and the DependencyParseAnnotator parsers,[27] which are the components of the Stanford CoreNLP package, the RelEx parser (Fundel, Küffner/Zimmer 2007) as well as the MST parser.[28]

6.2.2 Processing at Semantic, Pragmatic and Discourse Levels

6.2.2.1 Semantic Role Labeling

Semantic Role Labeling (SRL), or shallow semantic parsing is a process of automatically determining and assigning semantic labels, or thematic roles (also theta-roles, semantic roles), to the noun phrases. Thematic roles as defined Löbner (2013) are "different arguments of a verb predicate". A universal set of thematic roles, including a short description and an example, is provided in Table 24. For example, given a sentence *The child opened the door with her own key* (Löbner 2013: 122), the task of SRL would be to automatically recognize *child* as an animate agent, *door* as theme, and *key* as an instrument.

Role	Description	Examples
Agent/Actor	Performs the action expressed by the verb, controls the event	*Johnny* wrote a love letter.
Theme/Patient	Undergoes the action/change/event expressed by the verb	Johnny wrote *a love letter.*
Experiencer	Experiences a perception, feeling or other state	*I* heard him. The outburst surprized *her.*
Instrument	An instrument, or a cause, by which the event comes about	*This key* opens the door. He opened the door *with a key.*
Locative	A location	The keys are *on the desk.*
Goal	Goal of a movement	Put the keys *on the desk.*
Path	Path of a movement	She rode through *the desert.*

Table 24: Universal thematic roles. *Note*: From Understanding semantics (Löbner 2013: 123).

The early work on SRL (Hirst 1987; Richardson, Dolan/Vanderwende 1998) relied on the manual creation of semantic rule sets, which, however, was time-, work-, and cost-consuming and was usually restricted to a particular domain or a number of considered rules. In the 1990s, the application of machine learning methods in computational linguistics enabled systems to automatically acquire linguistic knowledge, rather than constructing it manually,

[26] http://www.maltparser.org/
[27] https://nlp.stanford.edu/software/stanford-dependencies.html
[28] http://www.seas.upenn.edu/~strctlrn/MSTParser/MSTParser.html

which, in turn, showed promising results for automatic semantic interpretation (Briscoe/Carroll 1997). Since then, a number of datasets of manually annotated text data with semantic roles have been presented, such as FrameNet (Fillmore, Ruppenhofer/Baker 2004), PropBank (Palmer, Gildea/Kingsbury 2005), VerbNet (Kipper et al. 2000), and NomBank (Meyers et al. 2004), which made it possible to develop statistical approaches specifically for the task of SRL.

FrameNet (Ruppenhofer et al. 2010) is considered to be the first corpus annotated with semantic roles, developed by Charles Fillmore at Berkeley's International Computer Science Institute. The main aim of the project was to represent frame semantics (Fillmore 1976) in a human and machine-readable database. The database is represented by a set of semantic frames. Each semantic frame includes a definition of the word meaning along with an example, a number of frame elements which represent the semantic roles within the frame, frame-frame relations establishing a relation to other frames in the database, and lexical units which evoke the frame. For example, a semantic frame for commerce_sell is composed, among others, of frame elements such as buyer, seller, item, money, place, and reason. The frame-frame relations for the frame commerce_sell are Inherits from: Giving is inherited by Renting_out, Perspective on Commerce_goods-transfer, etc. The lexical units that evoke the frame commerce_sell include, among others, auction.v, retail.v, vend.v. It should be noted that besides verbs, the corpus also includes semantic role annotation of nouns, adjectives, adverbs, and prepositions. Lexemes that belong to different parts of speech (*to construct* and *construction*) can be assigned to the same frame. The current version of the FrameNet for English contains 1224 lexical and non-lexical frames with 10542 frame elements in lexical frames and 13639 lexical units. More statistics on the current status of the project can be found on the website of the project.[29] The analogous FrameNet databases are currently available for a number of languages, including, among others, Chinese, Japanese, Korean, French, Swedish, Spanish, etc. FrameNet is an open-source resource which can freely be downloaded and can also be used on the web.

The Propositional Bank, or PropBank (Kingsbury et al. 2002; Palmer et al. 2005), is a corpus of annotated semantic roles developed at the University of Pennsylvania 2001 within the Propositional Bank project by Martha Palmer and Paul Kingsbury. The initial aim of the project was to add a semantic layer to the syntactic trees of the Penn Treebank (Marcus et al. 1993) by means of annotating predicates and the semantic roles of their arguments. The arguments of each verb are provided by a number such as, e.g. in (6.1), where arg0 generally designates a prototypical agent and arg1 marks a prototypical patient or theme (for

[29] https://framenet.icsi.berkeley.edu/fndrupal/

more information on prototypical roles, see Dowty 1991). Other arguments include the arg2 instrument/attribute, the arg3 starting point/attribute, the arg4 ending point, and the argM modifier. It should be noted that arguments with higher numbers are not consistently generalized.

(6.1) John broke the window: *broke(arg0 = John, arg1 = the window)*

The window broke: *broke(arg1 = the window)*

The PropBank consists of frame files where each frame file contains one or more verb senses. Each sense of a polysemous verb is referred to as a frameset. Each frameset is provided with a separate role set, which is a set of numbered roles. Further, each verb sense in a frameset is annotated with semantic roles as well as examples and links to other lexical tools such as FrameNet (described above) and VerbNet (described below). For example, the frame file for the verb to break consists of nine framesets, and the first frameset contains four roles, arg0 breaker, arg1 thing broken, arg2 instrument, and arg3 pieces. In total, the PropBank provides the semantic annotation of 1 million words, including 3,633 verbs. The corpus is an XML-annotated dataset. The PropBank open-source; the download link, as well as further information on the resource, can be found on the website of the project.[30]

The NomBank (Noun Annotation Bank) database (Meyers et al. 2004) is a resource developed at New York University and is based on the Nomlex project. The main aim of the NomBank project was to extend the PropBank corpus with the annotations of nouns. The developers define this aim as “to mark the sets of arguments that co-occur with nouns just as PropBank records such information for verbs”. Currently, the NomBank consists of the data of the Nomlex project as well as frame files for all “markable” nouns of the PropBank corpus. The data in the resource is annotated using XML. The current version of NomBank 1.0 consists of 114,576 propositions in total. The resource is freely available for download from the website of the NomBank project in .zip and .tgz file formats.[31]

VerbNet (Kipper et al. 2000) is a lexicon of English verbs developed by Karin Kipper-Schuler at the University of Pennsylvania. The verbs in the lexicon are grouped in accordance with their syntactic behavior, following Levin’s classification of verbs (Levin 1993). The current VerbNet version 3.2 contains about 4,000 verbs and 283 verb classes with 23 thematic roles. Each verb class contains a set of verbs based upon their meaning (members), thematic roles, and selectional restrictions on the arguments, as well as frames, including, besides an example, also a syntactic and semantic description. While a syntactic description

[30] http://verbs.colorado.edu/~mpalmer/projects/ace.html
[31] http://nlp.cs.nyu.edu/meyers/NomBank.html

provides all the possible realizations of the verb in the class and consists of a verb and the thematic roles of the arguments around the verb, semantic description lists semantic predicates. A simplified entry for the class *hit-18.1* is illustrated in Table 25.

Class hit-18.1			
Roles and Restrictions: Agent[+int_control] Patient[+concrete] Instrument[+concrete]			
Members: bang, bash, hit, kick, ...			
Frames:			
Name	Example	Syntax	Semantics
Basic Transitive	Paula hit the ball	Agent V Patient	cause(Agent, E)manner(during(E), directed-motion, Agent) contact(during(E), Agent, Patient) manner(end(E), forceful, Agent) contact(end(E), Agent, Patient)

Table 25: Simplified VerbNet entry for the *hit-18.1* class.

The VerbNet lexicon is an XML-annotated resource. It is open-source and can freely be downloaded from the website of the Department of Linguistics at the University of Colorado at Boulder.[32]

The first system for SRL has been proposed by Gildea and Jurafsky (2000) and was based on the FrameNet. Since the work of Gildea and Jurafsky (2000), a number of other systems have been developed. These have been presented within the seven international evaluation tasks in ACL-related conferences and workshops, the SIGNLL CoNLL shared tasks, which took place in 2004 and 2005 (Carreras/Màrquez 2004, 2005), the SIGLEX Senseval-3 in 2004 (Litkowski 2004) as well as four tasks in the SIGLEX SemEval in 2007 (Pradhan et al. 2007; Màrquez et al. 2007; Baker, Ellsworth/Erk 2007). To achieve a comparability of the results, the systems were evaluated on a common dataset. That is, the systems submitted by the Senseval-3 were evaluated on the FrameNet, while CoNLL-2004, CoNLL-2005, and SemEval 2007 used PropBank for the evaluation. The best result, e.g. on the CoNLL-2004, was shown by the system developed by Hacioglu et al. (2004), with an F-measure of 69.49%; on the CoNLL-2005, the best performance was shown by the systems proposed by Koomen et al. (2005) and Haghighi et al. (2005).

In general, two approaches to SRL have been followed so far, including probabilistic and machine learning, while application of machine learning algorithms is clearly a preferred approach. At present, the state-of-the-art SRL systems include the Illinois Semantic Role Labeler (Punyakanok, Roth/Yih 2008) and the Shalmaneser (Erk/Padó 2006).

[32] http://verbs.colorado.edu/verbnet_downloads/downloads.html

The Illinois Semantic Role Labeler[33] (Punyakanok et al. 2008) is a machine learning-based tool that identifies the thematic roles of noun phrases such as the agent, the patient, and the theme in a given sentence. It assigns labels to the noun phrases, following the notation defined by the PropBank project. The methodology of the system combines a machine-learning algorithm with inferencing that is based on an integer linear programming, incorporating linguistic and structural constraints into the decision process. Additionally, the scientist showed that full syntactic parsing information is essential for identifying an argument, especially in the first stage of the analysis. The system has been evaluated on all development and test datasets provided in the CoNLL-2005 shared task on semantic role labeling, including CoNLL-2005 development dataset, the Wall Street Journal test corpus, the Brown test corpus, and the last two mentioned corpora combined, and could achieve the highest F1 score (77.35%, 79.44%, 67.75%, and 77.92%, respectively) among 19 participants. The current version of the tool is available as a component of the Illinois NLP Curator.[34]

The Shalmaneser (Erk/Padó 2006) is a supervised learning SRL toolbox which runs under Linux and requires the Collins parser, TnT, and TreeTagger installed for applying the Shalmaneser[35] on English texts. The tool was developed for frame semantics and uses frame semantics terminology. For the end user, the Shalmaneser is provided in a simple user mode for applying pre-trained classifiers for English (trained on the FrameNet) and for German (trained on SALSA Frame annotation); it can, however, be further extended if required. The tool can be freely obtained by sending an email to the developer.

6.2.2.2 Recognizing Textual Entailment

The goal of the Recognizing Textual Entailment task, or RTE, is to determine "whether one piece of text can be plausibly inferred from another" (Dagan et al. 2009). That is, a textual entailment is observed if an ordinary speaker with a basic world and linguistic knowledge infers a hypothesis from a text, in terms of an RTE challenge, it is formulated as whether the text (*T*) entails the hypothesis (*H*) or not. Thus, the term textual entailment in NLP is used in a manner looser than logical entailment. Compare it to the classical "strict" definition of logical entailment provided in Chierchia and McConnell-Ginet (2001: 19-20):

> "A entails B (a) whenever A is true, B is true, (b) the information that B conveys is contained in the information that A conveys, (c) a situation describable by A must also be a situation describable by B and (d) A and not B is contradictory (can't be true in any situation)."

[33] http://cogcomp.org/page/software_view/SRL
[34] http://cogcomp.org/page/software_view/Curator
[35] http://www.coli.uni-saarland.de/projects/salsa/shal/

In comparison to the definition of textual entailment, the classical understanding of logical entailment does not account for some uncertainty in natural language applications such as in (6.2).

(6.2) T: The technological triumph known as GPS was incubated in the mind of Ivan Getting. (Agerri 2008: 3-4)

H: Ivan Getting invented the GPS.

A number of RTE approaches have been proposed till now. These include, among others, the application of machine learning techniques, with SVM preferred, the modeling of logical inference, the computation of the cross-pair similarity between T and H as well as the word alignment (for more elaboration on the diverse approaches to the RTE task, see Dagan et al. 2013). In order to provide a common basis for the evaluation of systems, a series of RTE challenges has been organized. The overview and summarization of the proposed methods and systems within the RTE challenge, as well as description of evaluation data, are provided in Dagan et al. (2006) for RTE-1, Bar-Haim et al. (2006) for RTE-2, in Giampiccolo et al. (2007) for RTE-3, in Giampiccolo et al. (2008) for RTE-4, Bentivogli et al. (2009) for RTE-5, Bentivogli et al. (2010) for RTE-6, and Bentivogli et al. (2011) for RTE-7.

At present, the state-of-the-art systems for performing RTE tasks are represented by VENSES, EXCITEMENT, EDITS, and Nutcracker.

VENSES (Venice Semantic Evaluation System)[36] is a system for performing the task of textual entailment recognition, developed by Rodolfo Delmonte (e.g. Delmonte 2009). The system is based on two subsystems: GETARUN and a Semantic Evaluator which was previously created for the Summary and Question evaluation. GETARUN is a system for text understanding which interprets meanings from complete linguistic representations. It consists of two processing modules. A low-level processing module includes a tokenizer, a tagger, a parser, an interpretation for grammatical relations and semantic roles assignment, quantifier raising, and pronominal binding. In turn, a high-level processing module includes a discourse model, a centering-like algorithm, a temporal interpretation, a logical form production, and a discourse structure analysis. The output of the system is a flat list of head-dependent structures, including syntactic (grammatical) relations as well as semantic roles, modality, mood, and negation relation information. The Semantic Evaluator is independent of the GETARUN and consists of two modules: a sequence of linguistic rule-based submodules, which are a sequence of syntactic-semantic transformation rules, and a quantitatively based measurement of input structures. The latter conducts a count of heads,

[36] http://project.cgm.unive.it/venses.html

dependents, and the number of grammatical and semantic relations, scoring only similar elements in the text-hypothesis pair. VENSES is an open-source program which can be used on Mac OS X 10.4.11, Mac OS X 10.5.5, as well as on Windows XP or later and Ubuntu Linux. Using VENSES requires the installation of SWI Prolog.

The EXCITEMENT Open Platform (EOP)[37] is an open-source platform for textual inferencing (involving the RTE task) for multiple languages developed as a part of the EXCITEMENT (EXploring Customer Interactions through Textual EntailMENT) project (Magnini et al. 2014; Padó et al. 2015; Zanoli/Colombo 2016). The current EOP version can be applied for English, German, and Italian and includes special tools for creating, training, and evaluating new resources in/for other languages. The platform includes a pipeline for conducting a linguistic pre-processing (e.g., tokenization, POS tagging, parsing) based on the Apache UIMA framework, the state-of-art algorithms (e.g. Entailment Decision Algorithm (EDA)), a large number of knowledge resources (e.g. WordNet, Wikipedia), and techniques for experimenting and testing innovative approaches. The RTE system BIUTEE (Bar Ilan University Textual Entailment Engine) (Stern/Dagan 2012) developed earlier is a part of the EXCITEMENT project as well. The EOP platform is an open-source tool and is released under General Public License (GPL) Version 3.

EDITS (Edit Distance Textual Entailment Suite)[38] is an another open-source RTE tool developed by Kouleykov and Magnini and tested by Negri, Cabrio, and Mehdad (Kouleykov/Negri 2010; Mehdad/Magnini 2009; Mehdad 2009). It recognizes the entailment relation between T and H by using the edit distance algorithm and computing the *T-H* distance as the cost of the edit operations (i.e. insertion, deletion, and substitution) that are necessary to transform T into H. The tool consists of three main modules: the edit distance algorithm, a cost scheme for three edit operations, and a set of rules expressing either entailment or contradiction. The main advantage of the tool is that its modules and parameters can be easily configured by the users. Moreover, the tool allows the processing of data at different levels of complexity. At present, EDITS can be applied to either Italian or English. The current version of EDITS is 2.1 which is freely available on its website.

Nutcracker is one of the first RTE systems which conducts the RTE task by means of inferencing and is based on FOL and the theorem prover developed by Bos and Markert (Bos/Markert 2005). The methodology underlying the system can be described as follows (van Harmelen, Lifschitz/Porter 2008: 806ff). First, the text and hypothesis are represented by discourse representation structures and then by the FOL formulas *T* and *C*, respectively.

[37] http://hltfbk.github.io/Excitement-Open-Platform/
[38] http://edits.fbk.eu/

In the second step, a relevant world knowledge, if required, is identified and presented by the FOL formula *BK*. Finally, an automated reasoning by means of the FOL theorem prover or model builder is performed in order to conclude whether an entailment relation $T \wedge BK \rightarrow C$ is observed. The world knowledge is generated by two means, including the WordNet as well as syntax and lexical information provided by the parser. Though the Nutcracker has been widely used by the linguistic community, it is not currently available for use.

6.2.2.3 Anaphora Resolution

6.2.2.3.1 Cohesion and Coherence

It is doubtless that sentences in a text are not just loosely lined up but rather connected with each other in a certain manner following certain principles and mechanisms. In our opinion, the knowledge of how this is realized can have a positive impact on modeling and performing an automatic detection of contradictions occurring in news texts.

Different attempts have been made to identify the mechanisms according to which the sentences are bound together in a text. The most prominent one includes the theory of de Beaugrande and Dressler (1981). Other catalogs of textuality have been proposed in Heinemann and Heinemann (2002), Schröder (2003), and Stede (2007).

De Beaugrande and Dressler (1981: 7) identify seven features of textuality, including two text-centered features, “designating operations directed at the text materials”, which are *cohesion* and *coherence* and five user-centered features “which are brought to bear on the activity of textual communication at large, both by producer and by receivers”, including intentionality, acceptability, informativity, situationality, and intertextuality, while cohesion and coherence enjoy the most attention in the research community. De Beaugrande and Dressler (1981:3) define cohesion as “the ways in which components of the *surface text*, i.e. the actual words we hear or see, are *mutually connected within a sequence*”. Coherence in turn “concerns the ways in which the components of the textual world, i.e. the concepts and relations which underlie the surface text, are mutually accessible and relevant” (p. 4). In other words, while cohesion refers to the connection between the sentences of a text, coherence refers to the connection of ideas conveyed the text.

The prominent theories on relational coherence include, among others, the theory proposed in Hobbs (1985) as well as the Rhetorical Structure Theory (RST) developed by Mann and Thompson (1988), Taboada and Mann (2006), and the theories described in Sanders, Spooren and Nordman (1992), and Kehler (2002).

Due to the text surface orientation of the present study on automatic CD, realization devices of cohesion are of particular interest. The prominent work on cohesion devices, or cohesion

ties, is considered to be the work of Halliday and Hasan (1976). Other similar catalogs of cohesion ties are proposed in Linke et al. (1994) and Bussmann (2006).

Halliday and Hasan (1976) distinguish between two types of cohesion ties – grammatical and lexical. Scientists count reference, substitution, ellipsis, and conjunction to the grammatical cohesion ties and reiteration and collocation to the lexical cohesion ties.

Reference is designated to the items which, "instead of being interpreted semantically in their own right, (…) make reference to something else for their interpretation. In English, these items are personals, demonstratives and comparatives" (Halliday/Hasan 1976: 31). In other words, a reader can decide what is being talked about by referring to another expression in the text. It is to be noted that the notion of reference in the understanding of Halliday and Hasan (1976) means the relation between the "expressions in different parts of a text", which is to be distinguished from the traditional view on reference which is a relation that "holds between expressions in a text and entities in the world" (Brown/Yule 1991: 204). In order to avoid terminological confusion, Brown and Yule (1991) propose to use the term co-reference when addressing reference in the sense of Halliday and Hasan.

According to Halliday and Hasan (1976), reference can be of two types, endophoric and exophoric reference. While the interpretation of endophoric items lies within a text, the interpretation of exophoric items occurs outside the text in the context of the situation. Further, Halliday and Hasan divide endophoric reference items into cataphoric (forward-pointing) and anaphoric (backward-pointing) such as exemplified in (6.3a) and (6.3b), respectively (examples are taken from Brown/Yule 1991: 193). Anaphoric reference, in turn, is regarded as the only relevant one for cohesion as it "provides a link with a preceding portion of the text" (Halliday/Hasan 1976: 51).

(6.3) a. It's going down quickly, the sun. (It refers forward to the sun)

b. Look at the sun. It's going down quickly. (It refers back to the sun)

From a functional point of view, the (cohesion) reference can be of three types, personal, demonstrative, and comparative. While the personal reference keeps track of information through persons by means of personal pronouns (e.g., *I*, *he*, *she*) and possessive determiners (e.g., *mine*, *yours*, *his*, *hers*) and demonstrative reference through location by proximity references (e.g., *this*, *these*, *that*, *those*, *here*, *there*, *then*, and *the*), comparative reference fulfils this function by setting up the relation of the identity and similarity between referents using adjectives (e.g., *identical*, *same*, *equal*, *similar*, *different*, *else*, *better*, *more*), adjectives in a comparative form (e.g., *bigger*, *faster*), and adverbs (e.g., *identically*, *likewise*, *so*, *such*, *similarly*, *otherwise*, *more*).

Two grammatical cohesion ties similar to reference are substitution and ellipsis. While (cohesion) reference operates at the semantic level, substitution and ellipsis establish links at the lexicogrammatical level. Substitution and ellipsis are used when "a speaker or writer wishes to avoid the repetition of a lexical item and draw on one of the grammatical resources of the language to replace the item" (Bloor/Bloor 1995: 96).

Substitution is a grammatical relation between words or phrases, referring to a replacement of one item by another. As already mentioned, the main difference between reference and substitution lies in the observation that "substitution is a relation in the wording rather than in the meaning" (Halliday/Hasan 1976: 88). Halliday and Hasan (1976) distinguish between three ways of substituting in English, including nominal substitution, which substitutes nouns by means of *one*, *ones*, *the same* such as in (6.4a), verbal substitution, substituting verb typically by using the verb *do* (often combined with *so*) as exemplified in (6.4b) and clausal substitution, substituting an entire clause usually by *so*, *not* as in (6.4c). Consider the following examples taken from Halliday and Hasan (1976: 89-90):

(6.4) a. My axe is too blunt. I must get a sharper one.

b. You think Joan already knows? – I think everybody does.

c. Has Barbara left? – I think so.

In turn, an ellipsis (or zero substitution) is a deletion of words, expressions, or phrases such as in (6.5a, b, c). The idea underlying ellipsis can be described as "something left unsaid". Similar to substitution, the types of ellipses also include the nominal (omission of a head of a noun phrase), the verbal (omission of the lexical verb), and the clausal (ellipsis of a large part of clauses) ellipsis. These are exemplified in (6.5a), (6.5b), and (6.5c), respectively (the following examples are taken from Halliday and Hasan (1976: 148ff.)):

(6.5) a. Four other Oysters followed them, and yet another four.

b. Have you been swimming? – Yes, I have.

c. What were they doing? – Holding hands.

Finally, conjunction as a type of grammatical cohesion tie includes specific devices for linking clauses or portions of text with each other, demonstrating a relation between them. Halliday and Hasan (1976) distinguish between additive, adversative, causal, and temporal types of conjunction. The typical means of additive conjunction include *and*, *also*, *too*, *furthermore*, *additionally*, as well as *nor*, *and…not*, *either*, *neither*, indicating a coordination or an addition relation between the clauses. Adversative conjunctions, in turn, express a "contrary to expectation" relation and are realized by *yet*, *though*, *only*, *but*, *in fact*, *rather*, etc. Causal conjunctions indicate result, reason, and purpose and are realized by *so*, *then*, *for*, *because*, *for this reason*, *as a result*, *in this respect*, etc. Finally, temporal conjunctions

indicate a sequence or temporal relation and are signaled by *then*, *next*, *after that*, *next day*, *until then*, *at the same time*, *at this point*, etc.

In contrast to grammatical cohesion, lexical cohesion is not grammatical and refers to the "cohesive effect achieved by the selection of vocabulary" (Halliday and Hasan 1976: 274). Halliday and Hasan (1976) distinguish between two types of lexical cohesion, reiteration and collocation. Reiteration, which "involves the repetition of a lexical item, at one end of the scale" (Halliday/Hasan 1976: 278) can be realized through repetition, synonymy, hyponymy (also subordinate) (6.6), metonymy and antonymy.

(6.6) Henry's bought himself a new *Jaguar*. He practically lives in the car (subordinate).

Collocation in turn refers to "co-occurrence of lexical items that are in some way or other typically associated with one another because they tend to occur in similar environments" such as *boy…girl*, *stand up…sit down*, *car…brake*, *garden…dig*, etc. (Halliday and Hasan 1976: 287). Further, Morris and Hirst (1991: 22-23) point out that "often, lexical cohesion occurs not simply between pairs of words but over a succession of a number of nearby related words spanning a topical unit of the text. These sequences of related words will be called *lexical chains*. (…) *Lexical chains* tend to delineate portions of text that have a strong unity of meaning." Lexical chains for the excerpt in (6.7) are *mountaineering…Yosemite…summit peaks…climb…ridge* (L1) and *riding…ride…travel* (L2).

(6.7) Few Yosemite visitors ever see snow avalanches and fewer still know the exhilaration of riding on them. In all my mountaineering I have enjoyed only one avalanche ride, and the start was so sudden and the end came so soon I had but little time to think of the danger that attends this sort of travel, though at such times one thinks fast. One fine Yosemite morning after a heavy snowfall, being eager to see as many avalanches as possible and wide views of the forest and summit peaks in their new white robes before the sunshine had time to change them, I set out early to climb by a side canyon to the top of a commanding ridge a little over three thousand feet above the Valley. (Example is taken from Halliday/Hasan 1976: 286).

6.2.2.3.2 Anaphora Resolution: Approaches and Tools

The task of finding the reference of a noun phrase is referred to in NLP as anaphora, or coreference resolution. The first systems for anaphora resolution appeared in the 1960s, an intensive research on the task, however, began in the 1990s (Mitkov 2002). At present, a number of methods exist for the task (Table 26). According to the underlying methodology, these can be divided into two groups – rule-based and data-based.

A rule-based approach – an early approach to anaphora resolution – is based on the heuristics about the phenomenon of anaphora. Three kinds of rule-based methods have been proposed up to now, which are syntax-based, discourse-based, and hybrid anaphora resolution. While a syntax-based anaphora resolution relies on syntactic rules, making use of parsing, the discourse-based methods (e.g. Kibble 2001, Tetreaut 2001) are built on

discourse theories such as Centering Theory (Grosz et al. 1995). The first syntax-based algorithm which showed a high performance (an accuracy of 88.3%) is considered to be the Hobb's naïve algorithm (Hobbs 1976). An example of a hybrid approach is represented, e.g., by Lappin and Leass (1994). Existing rule-based methods are summarized in Table 26 (second column). In general, the rule-based approach can be characterized as labor-intensive as the methods are evaluated manually.

Type of cohesion tie	Rule-based approaches	Data-based approaches
Central pronouns	Hobbs (1976), Brennan, Friedman and Pollard (1987)*, Lappin and Leass (1994), Kennedy and Boguraev (1996), Baldwin (1997), Hobbs and Kehler (1997), Mitkov (1998)* and (2002)*, Stuckhardt (2001), Tetreault (2001)*, Byron (2002), Haghighi and Klein (2009)	Evans (2001), Soon, Ng and Lim (2001), Ng and Cardie (2002a), Ng and Cardie (2002b)*, Poesio and Kabadjov (2004), Boyd, Gegg-Harrison and Byron (2005), Versley et al. (2008)*, Stoyanov et al. (2010)*, Uryupina (2010)*
Reciprocal pronouns	Lappin and Leass (1994), Kennedy and Boguraev (1996), Stuckhardt (2001)	-
Demonstrative pronouns	Byron (2002)	Soon, Ng and Lim (2001), Ng and Cardie (2002a), Ng and Cardie (2002b)*, Versley et al. (2008)*, Stoyanov et al. (2010)*, Uryupina (2010)*, Kolhatkar and Hirst (2012)
Relative pronouns	Stuckhardt (2001)	Cardie (1992)
Adverbs	-	-
Noun phrases with a definite article	Bean and Riloff (1999), Vieira and Poesio (2000), Stuckhardt (2001), Meyer and Dale (2002), Markert and Nissim (2005), Haghighi and Klein (2009)	McCarthy and Lehnert (1995), Vieira and Poesio (2000), Soon, Ng and Lim (2001), Ng and Cardie (2002a), Ng and Cardie (2002b)*, Poesio and Kabadjov (2004), Versley et al. (2008)*, Stoyanov et al. (2010)*, Uryupina (2010)*
Proper names	Stuckhardt (2001), Haghighi and Klein (2009)	Soon, Ng and Lim (2001), Ng and Cardie (2002a), Ng and Cardie (2002b)*, Poesio and Kabadjov (2004), Versley et al. (2008)*, Stoyanov et al. (2010)*, Uryupina (2010)*
Indefinite pronouns	Markert and Nissim (2005)	Ng et al. (2005)
Other forms of coreference and substitution	Markert and Nissim (2005)	-
Verb phrases with do and combinations	Hardt (1997), Hobbs and Kehler (1997), Asher, Hardt and Busquets (2001), Kehler (2002)	Nielsen (2004)
Ellipses	Hardt (1997), Hobbs and Kehler (1997), Asher, Hardt and Busquets (2001), Kehler (2002)	Nielsen (2004)
Non-finite clauses	-	-

Table 26: Overview of anaphora resolution methods (*Types of cohesion tie are not treated in detail). *Note*: From Anaphora resolution and text retrieval (Schmolz 2015: 236).

While rule-based methods rely on the linguistic knowledge of anaphora, data-based (machine learning based) methods use data with manually annotated (or labeled) anaphoric relations and apply machine learning techniques. From the annotated data, the machine learning algorithms "learn" the rules which are then applied to anaphora resolution in unannotated data. The preferred machine learning algorithms for anaphora resolution are decision trees and instance-based methods. The first successful data-based methods for anaphora resolution involve McCarthy and Lehnert (1995) and Soon, Ng and Lim (2001) for English and Aone and Bennett (1995) for Japanese. Soon, Ng, and Lim (2001), for instance, evaluated their method on the training and test corpora of MUC-6 and MUC-7 with a recall of 58.6%, a precision of 67.3%, and an F-measure of 62.6%. According to Schmolz (2015: 233) who followed Mitkov (2002:113), Mooney (2004: 376-377), Schmid (2010: 180-181), and Strube (2010: 400-407), the main advantages of the data-based approach are their domain- and language independence, as well as their robustness in comparison to the rule-based approach and an ability of noticing connections of factors which a human might overlook. The data-based approaches are summarized in Table 26 (third column).

Among the off-the-shelf tools that can be applied for automatic anaphora resolution are BART, IMSCoref, HOTCoref as well as the coreference resolution components of the Stanford CoreNLP, only to name a few.

BART[39] (earlier: Baltimore Anaphora Resolution Toolkit, currently: Beautiful Anaphora Resolution Toolkit) is an off-the-shelf tool for the automatic processing of anaphora resolution based on machine learning. It was developed within the project Exploiting Lexical and Encyclopedic Resources for Entity Disambiguation at the Johns Hopkins Summer Workshop 2007 (Versley et al. 2008; Versley/Björkelund 2016). Primarily developed for English, BART can be successfully applied to other languages. It is provided with two modes – a web demo/trial mode and a mode for training and testing one's own data. One of the main advantages of the tool is that it is modular, incorporating the essential steps needed for the preprocessing of the data before the actual anaphora resolution can take place. Additionally, it allows the use of other machine learning tools such as WEKA and MaxEnt. To upload data into BART in plain text format, a REST-based web service can be used, which is a part of BART. Internally, BART works with a representation based on the format of the MMAX2 tool, which is an annotation tool for coreference and other discourse properties. It also supports the XML input format and delivers its output in XML. It also provides support for error analysis using the MMAX2 annotation tool. According to Versley and Björkelund (2016),

[39] http://www.bart-coref.org/

BART has been successfully applied to different corpora. Bart is open-source and is licensed under the Apache license (v2.0), except for a small number of classes.

IMSCoref is another tool for automatic anaphora resolution. It is a pure coreference resolver that does not include a preprocessing pipeline. The tool was developed by Björkelund and Farkas (2012) for the CoNLL-2012 shared task (Pradhan et al. 2012) at the University of Stuttgart and showed the second best overall result among the tools submitted. Since then, it has been improved by Björkelund and Kuhn (2012) for the COLING-2012 conference. The approach underlying the system (both versions) is a mention-pair model (Aone/Bennett 1995; Soon, Ng/Lim 2001; Ng/Cardie 2002b) which "recasts the coreference resolution problem as a classification task in which a classifier is trained to decide for a given pair of noun phrases whether they corefer or not. In a second step, full coreference chains are built by clustering these pairwise decisions" (Hoste 2016: 269). The main difference between the IMSCoref tool and the other tools which are also built on the pair-mention model is its configurable feature set. For the CoNLL version of the tool, the features have been set individually for each of the three languages Arabic, Chinese, and English. The second version of the tool additionally incorporates modifications in order to deal with dependencies. Both versions of the system are freely available for download at the website of the Institute for Natural Language Processing of the University of Stuttgart and are licensed under GNU General Public License.

HOTCoref[40] (Higher Order Tree Coreference) is an off-the-shelf tool for anaphora resolution (coreference resolution), also developed at the Institute for Natural Language Processing of the University of Stuttgart[41] by Björkelund and Kuhn (2014) and despite being built on the same codebase as IMSCoref, it is considered to be faster and more efficient than IMSCoref. Moreover, the tool shows the best performance among the tools submitted for the CoNLL-2012 shared task. HOTCoref is a data-driven tool which models the coreference of a document as a directed rooted tree. The tool is written in Java and is platform independent. The HOTCoref is an open-source tool and is licensed under the GNU General Public License.

Finally, the components of the Stanford CoreNLP should also be mentioned, which are the three coreference systems – deterministic, statistical and neural. The deterministic coreference system[42] is rule-based and can be applied on the data in English and Chinese (Raghunathan et al. 2010; Lee et al. 2011; Lee et al. 2013; Recasens, de Marneffe/Potts 2013). The system has been evaluated on the CoNLL-2012 evaluation data and showed an

[40] http://www.ims.uni-stuttgart.de/forschung/ressourcen/werkzeuge/IMSCoref.en.html
[41] http://www.ims.uni-stuttgart.de/forschung/ressourcen/werkzeuge/HOTCoref.en.html
[42] https://nlp.stanford.edu/software/dcoref.html

F1 score of 49.5% for English and 47.5% for Chinese. The statistical coreference annotator, in turn, is a machine learning system which, in contrast to other Stanford coreference systems, incorporates a faster dependency parser instead of a constituency parser (Clark/Manning 2015). The system is available only for English. It was also evaluated on the CoNLL-2012 evaluation data and showed a 56.2% F1 score. Finally, the neural coreference annotator is a neural network-based coreference resolution for English and Chinese (Clark/Manning 2016). It is the most accurate among the coreference resolvers mentioned, but it is slow. The computed F1 score achieved by the system evaluated on the CoNLL-2012 evaluation dataset amounts to 60.0% and 53.9% for English and Chinese, respectively.

6.2.3 Further Processing Tasks

6.2.3.1 Negation and Modality Processing

A doubtlessly essential component for the CD task is the automatic treatment of negation and modality which, in our opinion, has the greatest impact on the performance of the system.

Negation and modality processing is a new and challenging area of research in NLP (see Morante/Sporleder 2012 for more elaboration on negation and modality). The phenomenon of modality is usually treated in NLP in connection with subjectivity[43] (Banfield 2014; Wiebe et al. 2004, 2005; Wilson et al. 2006; Prabhakaran, Rambow/Diab 2010), hedging[44] (Lakoff 1973; Hyland 1998; Medlock/Briscoe 2007), evidentiality[45] (Aikhenvald 2004; von Fintel

[43] "Subjectivity is the language used to express private states in the context of a text or conversation. Private state is a general cover term for opinions, evaluations, emotions, and speculations." (Wieber et al. 2004).

[44] Lakoff (1973: 471) defines hedges as "words whose job it is to make things fuzzier or less fuzzy". Hedges refer to the words such as *sort of*, *largely*, *basically*, *especially*, *almost*, *literally*, *roughly*, etc.

[45] The definition of evidentiality provided by Aikhenvald (2004: 1) is as follows: "In a number of languages, the nature of the evidence on which a statement is based must be specified for every statement - whether the speaker saw it, or heard it, or inferred from indirect evidence, or learnt it from someone else. This grammatical category, referring to an information source, is called 'evidentiality'."

2006), uncertainty[46] (Rubin, Liddy/Kado 2006; Rubin 2007), and factuality[47] (Pustejovsky et al. 2003; Saurí/Pustejovsky 2007, 2009, 2012).

For the task of hedges scope resolution, two approaches have been preferred so far, including machine learning-based systems (e.g. Morante/Daelemans 2009a, 2009b; Zhu et al. 2010) and rule-based systems that mainly rely on syntactic information (Özgür/Radev 2009). For the tagging of modality string- and structure-based English, taggers have been developed (Baker et al. 2010). Belief categorization has been addressed, e.g., in Prabhakaran, Rambow, and Diab (2010). Machine learning and rule-based approaches have been applied in detecting speculative sentences in MEDLINE abstracts (Light, Qiu/Srinivasan 2004), biological articles FlyBase (Medlock/Briscoe 2007; Szarvas 2008), in Wikipedia (Ganter/Strube 2009), and other datasets (Kilicoglu/Bergler 2008).

Most of the work on automatic negation processing focuses on deciding whether a term is negated or not. For instance, rule and regular expression-based systems like NegEx (Chapman et al. 2001) and NegFinder (Mutalik et al. 2001) as well as machine learning systems (Averbuch et al. 2004) and hybrid systems which combined rule-based methodology with machine learning techniques have been proposed for this task, e.g., for negation detection in biomedical texts (e.g. Huang/Lowe 2007). The task of negation scope resolution in the BioScope corpus has been addressed in Morante and Daelemans (2009b) and for sentiment analysis, in the BioScope corpus and Google Product Reviews Corpus in Councill, McDonald, and Velikovich (2010).

Existing annotated resources for the development and evaluation of methods and systems for negation processing include a corpus of Conan Doyle stories (Morante, Schrauwen/Daelemans 2011) annotated with negation scope and cues similar to the BioScope corpus (Vincze et al. 2008). It is to be noted that the BioScope corpus is additionally annotated with linguistic cues that express speculative language and its scope. For modality, a corpus annotated with modality categories (modality type, including polarity, volition,

[46] "(…) certainty is viewed as a type of subjective information available in texts and a form of epistemic modality expressed through explicitly-coded linguistic means. Such devices as subjectivity expressions, epistemic comments, evidentials, reporting verbs, attitudinal adverbials, hedges, shields, approximators, understatements, tentatives, intensifiers, emphatics, boosters, and assertives, often overlap in their definitions, classifications, and lexical representations in English. In essence, they perform the same role for the purpose of this study. They explicitly signal presence of certainty information that covers a full continuum of writer's confidence, ranging from uncertain possibility and withholding full commitment to statements to a confident necessity, reassurance, and emphasizing of the full commitment to statements." (Rubin et al. 2006: 65).

[47] According to Saurí and Pustejovsky (2009), factuality is defined as "information conveying whether events mentioned in text correspond to real situations in the world or, instead, to situations of uncertain status" as well as "The level of information expressing the commitment of relevant sources towards the factual nature of events mentioned in discourse".

obligation, belief, potential, permission, evaluative; scalar value ranging from zero to one; scope and attributed-to) in the framework of the OntoSem project (Nirenburg/Raskin 2004) are available, as well as FactBank (Saurí/Pustejovsky 2009) which is annotated with factuality information, including factuality degrees fact, counterfact, probable, not probable, possible, not certain, certain but unknown output, and unknown or uncommitted.

6.2.3.2 Sentiment Analysis

Sentiment analysis (also opinion mining, sentiment mining, subjectivity detection) can be defined as "the field of study that analyses people's opinions, sentiments, appraisals, attitudes, and emotions toward entities and their attributes expressed in written text. The entities can be products, services, organizations, individuals, events, issues, or topics." (Liu 2015: 1).

Sentiment analysis can be conducted on three levels of granularity, which are the document level, the sentence level, and the entity/aspect level (Liu 2015). Sentiment analysis on the document level, or document-level sentiment classification, classifies the whole document to positive, negative, or neutral (no opinion expressed) sentiment. That is, e.g., in a case of a product review, the system decides whether the review expresses an overall positive, negative, or neutral opinion about the product. Sentiment analysis on the document level can be applied on the text on one entity and is, therefore, not applicable to reviews on multiple entities. Sentiment analysis on the document level has been applied in, e.g., Pang, Lee, and Vaithyanathan (2002) to classify movie reviews, applying supervised machine learning.

Sentiment analysis on the sentence level, in turn, decides whether a sentence expresses a positive, negative, or neutral opinion and is closely related to the subjectivity classification (Wiebe et al. 1999). The subjectivity classification "distinguishes sentences that express factual information (called objective sentences) from sentences that express subjective views and opinions (called subjective sentences)" (Liu 2015: 9). However, as Liu further points out, subjectivity and sentiment/opinion cannot be regarded as synonymous. Many subjective sentences do not express any opinion or sentiment such as, e.g., in *I think he went home after lunch*. Also, many objective sentences can contain opinion and sentiment such as in *We bought the car last month and the windshield wiper has fallen off*. An example of applying a sentiment analysis on sentence level is described in Kudo and Matsumoto (2004).

Finally, sentiment analysis on the entity/aspect level (early also feature level) classifies the sentiment with respect to a certain entity or specific aspect of an entity. Thus, in the sentence, *Apple is doing very well in this poor economy*, the task of the aspect-level sentiment analysis would be to recognize that a negative sentiment is expressed towards the

economy, while the sentiment toward Apple's performance is positive. Del Pilar Salas-Zárate et al. (2017), for instance, examined tweets to apply an aspect-based sentiment analysis about diabetes.

At present, one distinguishes between three kinds of approaches to sentiment analysis, which are machine learning, lexicon-based techniques, and their hybrid approaches (D'Andrea et al. 2015).

Machine learning approaches apply machine learning algorithms and make use of data annotated not only with opinion words and phrases but also with term frequency, pos information, and negations. Bayesian networks, as well as naive Bayes, maximum entropy, neural networks, and SVM, belong to the preferred machine-learning algorithms for sentiment analysis. The main advantage of machine learning methods is an easy adaptation of trained models for specific purposes and contexts. However, a big amount of annotated data is needed, which is, in most cases, time- and cost-consuming.

In their turn, sentiment analysis methods that follow a lexicon-based approach rely on manually created lists of negative and positive terms and phrases. The lexicon-based approach can be further divided into dictionary-based, novel machine learning, corpus-based, and Ensemble approaches (for further elaboration, see D'Andrea et al. 2015). At present, a number of ready-to-use lexicons such as Sentiwordnet (Esuli/Sebastiani 2006) are provided. Though the lexicon-based approach is simple and does not require a large amount of annotated data, it is limited by the finite number of words in the lexicon.

Finally, the hybrid approach combines lexicon-based sentiment analysis with machine learning techniques. That is, the methods use a sentiment lexicon for identifying sentiment words, which are then used as features in applying the machine learning algorithms. An advantage of the hybrid sentiment analysis is its applicability at the aspect level as well as its low sensitivity to changes in the data domain. A problematic for the hybrid approach are noisy data.

The off-the-shelf tools for sentiment analysis include LIWC (Tausczik/Pennebaker 2010), SentiStrength (Thelwall et al. 2010), SenticNet (Cambria et al. (2010), and PANAS-t (Gonçalves, Benevenuto and Cha (2013), among others. Additionally, for conducting a sentiment analysis, an open source component of the Stanford CoreNLP – SentimentAnnotator can be applied,[48] which implements a sentiment model proposed in Socher et al. (2013). The preferred sources for sentiment analysis in Python are the TextBlob[49] and NLTK libraries.

[48] https://nlp.stanford.edu/sentiment/
[49] http://textblob.readthedocs.io/en/dev/

An NLTK sentiment analysis can be applied to the data in English, Dutch, and French. The English sentiment analysis uses classifiers (e.g. Naïve Bayes) trained on twitter data and movie reviews from the datasets[50] created on the NLTK-trainer[51] and described in Pang, Lee, and Vaithyanathan (2002). The output is a classification of the analyzed data into the positive, negative, or neutral sentiment. Hereby, neutrality is determined first, and the sentiment polarity is determined second in the case the text is not neutral. Besides NLTK, the sentiment analysis module (textblob.sentiments) included in the TextBlob 0.5.0 library can be used. The textblob.sentiments module contains two sentiment analysis implementations – the PatternAnalyzer which is based on the pattern library,[52] and the NaiveBayesAnalyzer which is an already mentioned NLTK classifier trained on a movie reviews corpus.

6.2.3.3 Named-Entity Recognition

Named-entity recognition, or NER, refers to the automatic identification of all the mentions of noun phrases, the so-called named entities (NE) occurring in a text or in a collection of texts and their classification into a set of predefined types. The universal types of NEs include PERSON (e.g. *Angela Merkel*), ORGANIZATION (e.g. *Union Investment*), and LOCATION (e.g. *Grand Canyon*). The common types of NEs are DATE (e.g. *24.01.2018*), TIME (e.g. *20:22 p.m.*), MONEY (e.g. *10,000 rubles*), PERCENT (e.g. *100%*), FACILITY – architectural artefacts (e.g. the *Pisa tower*) and GPE – geopolitical entities such as country, state, city, or town (e.g., *Giessen*). Additionally, NEs can be of domain-specific types such as names of drugs or bibliographic references. The types mentioned are not exhaustive and can vary for different NER systems and domains.

The task of an NER often provides a basis for an RTE task (Section 6.2.2.2) or an anaphora resolution (Section 6.2.2.3). The existing methods of NER can be classified into four groups of approaches, including list-based (or gazetteer-based), rule-based, machine learning-based, and hybrid approaches.

The use of lists (or gazetteer) of named entities is the widely used approach for automatically recognizing named entities. The main idea underlying this approach is to the recognize noun phrases occurring in textual data based on predefined lists. It is the simplest, fastest technique that is easy to retarget. One should, however, consider that the system recognizes only the entities stored in the lists. The process of collection and maintenance of lists, in turn, can be very labor-intensive. Moreover, the NEs are too numerous to include in dictionaries entirely, occurring with the problem that NEs of new types permanently appear.

---

[50] http://www.cs.cornell.edu/people/pabo/movie-review-data/
[51] https://github.com/japerk/nltk-trainer or here https://bitbucket.org/japerk/nltk-trainer/src
[52] https://www.clips.uantwerpen.be/pattern

For instance, Sekine et al. (2002) proposed an NER system that works with 150 categories. Further problematic issues related to the list-based NER are, on the one side, the inability of the method to deal with name variants, such as *John Smith*, *Mr. Smith*, and *John*, if these are not included in the list. On the other side is the difficulty to deal with the ambiguity of NE types such as *John Smith* (company vs. person), *May* (person vs. month), *Washington* (person vs. location), and *1945* (date vs. time). Although it is easy to develop and use, a list-based NER does not perform well.

The identification of rule-based NEs is realized by means of manually coded rules. For instance, a rule for extracting locations in English can be an occurrence of a capitalized word in combination with words such as *city*, *center*, *river*, *street*, *boulevard*, and *avenue*. This rule would, for instance, match *New York city* and *Main river*. Although capitalization is a strong indicator for proper names in English, some exceptions such as the capitalization of the first word of a sentence, as well as of those titles, have still to be considered. The rule-based NER is a very precise approach and requires a small amount of training data. However, it is expensive to develop the test cycle, is domain-dependent, and has to cope with some of the problematic issues partially mentioned above.

Machine learning-based NER depends on supervised and semi-supervised machine learning. The preferred algorithms for supervised learning include the Hidden Markov Models as well as the k-nearest neighbors, decision trees, AdaBoost (Carreras et al. 2002), and SVM. The main advantages of the machine learning NER annotators that they are easily applied to different languages as well as a higher recall rate than by rule- and list-based NER annotators. The main disadvantage of the approach is that it requires a large amount of manually labeled training data.

The ready annotated datasets for implementing the NER tools are represented, among others, by the corpora compiled in the framework of the CoNLL-2002 (Tjong Kim Sang 2002), CoNLL-2003 (Tjong Kim Sang/De Meulder 2003), MUC-6,[53] and MUC-7[54] conferences as well as by the ACE-2005 Multilingual Training Corpus[55] and BBN Pronoun Coreference and Entity Type Corpus.[56] Basic information on the corpora mentioned is summarized in Table 27.

[53] https://catalog.ldc.upenn.edu/LDC2003T13
[54] https://catalog.ldc.upenn.edu/LDC2001T02
[55] https://catalog.ldc.upenn.edu/ldc2006t06
[56] https://catalog.ldc.upenn.edu/ldc2005t33

Corpus	Data	Language	Types of NEs	Statistics	Availability
CoNLL-2002	Spanish EFE News Agency newswire, four editions of the Belgian newspaper *De Morgen*	Spanish, Dutch	Person, Location, Organization and Miscellaneous names	*Spanish dataset*: The data contains words and entity tags only. The training, development and test datasets contain 273037, 54837 and 53049 lines respectively. *Dutch dataset*: The data consists of words, entity tags and part-of-speech tags. The training, development and test datasets contain 218737, 40656 and 74189 lines respectively.	At the website of CLiPS (Computational Linguistics and Psycholinguistics Research Center)[57], freely available
CoNLL-2003	Reuters corpus (English) and Frankfurter Rundschau corpus from the ECI Multilingual Text Corpus (German)	English, German	Person, Location, Organization and Miscellaneous names	*English dataset*: Trainining set (203,621 tokens; 23,499 NEs) Development set (51,362 tokens; 5,942 NEs) Test set (46,435 tokens; 5,648 NEs) *German dataset*: Trainining set (206,931 tokens; 11,851 NEs) Development set (51,444 tokens; 4,833 NEs) Test set (51,943 tokens; 3,673 NEs)	At the website of CLiPS (Computational Linguistics and Psycholinguistics Research Center)[58], freely available
MUC 6	318 articles from the Wall Street Journal	English	Person, Location, Organization, Time, Date, Percent, Money	-	At the website of the LDC[59], requirement: membership or payment
MUC 7	158,000 articles from the New York Times	English	Person, Location, Organization, Time, Date,	-	At the website of the LDC[60], requirement: membership or payment

[57] https://www.clips.uantwerpen.be/conll2002/ner/
[58] https://www.clips.uantwerpen.be/conll2003/ner/
[59] https://catalog.ldc.upenn.edu/LDC2003T13
[60] https://catalog.ldc.upenn.edu/LDC2001T02

			Percent, Money		
ACE 2005	Weblogs, broadcast news, newsgroups, broadcast conversation	Mandarin Chinese, Standard Arabic, English	Location, Organization, Person, FAC, GPE	*Arabic dataset*: 150,000 words *Chinese dataset*: 300,000 words *English dataset*: 300,000 words	At the website of the LDC[61], requirement: membership or payment
BBN	Penn Treebank corpus of the Wall Street Journal texts	English	105 fine-grained tags: 54 corresponding to CONLL entities; 21 for numerical and time data; and 30 for other classes of terms	One million words	At the website of the LDC[62], requirement: membership or payment

Table 27: Available corpora with annotated named entities.

A preferred open-source tool for NER at present is the CRFClassifier component[63] of the Stanford CoreNLP. This tool is able to recognize the NEs of types PERSON, LOCATION, ORGANIZATION, and MISC as well as the numerical (MONEY, NUMBER, ORDINAL, PERCENT), and the temporal (DATE, TIME, DURATION, SET) entities in the data in German, English, Spanish, and Chinese. The CRFClassifier uses three Conditional Random Field (CRF) sequence taggers (Lafferty, McCallum/Pereira 2001; Finkel, Grenager/Manning 2005) trained for English on ACE, CoNLL, MUC-6, and MUC-7 datasets. Numerical entities are recognized using a rule-based system. For the normalization of numerical entities such as dates and measures, additionally, the NormalizedNamedEntityTagAnnotation component can be applied. The annotator is licensed under the GNU General Public License (v2 or later).

The Stanford CoreNLP also includes a TokensRegexNERAnnotator which conducts a simple, rule-based recognition of named entities over token sequences using regular expressions defined in Java. The main advantage of this NER annotator is that it provides a simple framework to incorporate non-conventional NE types such as IDEOLOGY, NATIONALITY, RELIGION, and TITLE. More information on the use of the TokensRegexNERAnnotator can be found on the website of the Stanford NLP Processing Group.[64]

[61] https://catalog.ldc.upenn.edu/ldc2006t06
[62] https://www.ldc.upenn.edu/
[63] https://nlp.stanford.edu/software/CRF-NER.html
[64] https://nlp.stanford.edu/software/regexner.html

6.2.3.4 Temporal Processing

Regarding parts of a contradiction that must have the same temporal reference (see Section 3.1.1), the automatic modeling of temporal information in a text is doubtless one of the essential tasks to have an impact on correct CD.

The task of temporal processing (also time processing) can be defined as “the automatic identification of all temporal referring expressions, events, and temporal relations within a text” (Marşic 2011: 2). Temporal expression (also time expressions), in turn, can be defined as “natural language phrases that refer directly to time, giving information about when something happened, how long something lasted, or how often something occurred” (Marşic 2011: 17). Temporal expressions in texts can occur as nouns (e.g., *week*, *past*, *year*), including proper names (e.g., *New Year*, *January*, *Monday*), adjectives (e.g., *next*, *daily*, *medieval*), adverbs (e.g., *currently*, *tomorrow*, *tonight*), conventional time patterns (e.g., *24/01/2018*, *80s*), and numbers (*3rd* as in *3rd October*).

Further, according to Biber et al. (1999) and Quirk et al. (2010), temporal expressions can refer to different kinds of time-related information such as position (e.g. I saw him *yesterday*), duration (e.g. He lived *several years* in Italy), frequency (e.g. He goes for a walk with his dog *every day*) and the relationship between events. Regarding the latter, Allen (1983, 1991), for instance, distinguishes between 13 relations, including the identity relations *before*, *meets*, *overlaps*, *starts*, *finishes*, *during*, *after*, *is met by*, *is overlapped by*, *is started by*, *is finished by*, and *contains*.

The task of temporal processing by machines faces a number of difficulties. The first one is caused by the observation that temporal information can be realized by a wide range of mechanisms such as aspect, tense, lexical semantic knowledge (Mani et al. 2006), only to name a few. Another difficulty arises as the result of that temporal information is not always expressed explicitly such as exemplified in (6.8). Despite the similar syntax in (6.8a) and (6.8b), events described in the sentences do not correspond to the same temporal order pattern. While in (6.8a), the second sentence (event: *pushed*) precedes the first sentence (event: *fell*), in (6.8b), the events take place in converse order. First, the event *fell* described in the first sentence takes place, and then the event *asked for help* occurs in the second sentence. Such cases are challenging for automatic temporal processing as no explicit time cues on the temporal anchoring of the events are provided here.

(6.8) a. John fell. Mary pushed him.
b. John fell. Mary asked for help. (Marşic 2011: 2)

As already mentioned above, the term temporal processing refers to four tasks, which are temporal expression identification, temporal expression normalization, event annotation,

and temporal relation identification. The key idea underlying the tasks can be illustrated by the following excerpt from a news article (Marşic 2011: 2).

> 27/02/1998
> OAU to investigate Rwandan genocide
> The Organization of African Unity said Friday it would investigate the Hutu-organized genocide of more than 500,000 minority Tutsis in Rwanda nearly four years ago. Foreign ministers of member-states meeting in the Ethiopian capital agreed to set up a seven-member panel to investigate who shot down Rwandan President Juvenal Habyarimana's plane on April 6, 1994.

That is, the aim of the task temporal expression identification would be to identify 'nearly four years ago' as a temporal expression. The task of temporal expression normalization would be to recognize that *nearly four years ago* refers to *1994*. The task of event annotation, in turn, would be to annotate the event *shot*, for example, as an occurrence. Finally, identifying the temporal relationship between the event *shot* and the temporal expression *April 6, 1994*, as an overlap is the key idea of temporal relation identification task.

Two approaches to temporal processing – machine learning (supervised and unsupervised) and rule-based – have been preferred so far. Up to now, a number of methods and systems for temporal processing have been proposed for the tasks of identifying and normalizing the temporal expressions[65] e.g. in Alexandersson et al. (1997), Busemann et al. (1997), Wiebe et al. (1998), Mani and Wilson (2000), Mazur and Dale (2008), and Chang and Manning (2012). The rule-based method proposed by Mani and Wilson (2000) for resolving temporal expressions in news texts showed an 83.2% accuracy, evaluated on their own dataset and manually annotated, and is considered to be widely used. Also worth mentioning are the LTG MUC system (Mikheev et al. 1998) and the Facile system (Ciravegna et al. 1999) submitted by the MUC conference series, the ATEL system (Hacioglu et al. 2005) submitted by the TERN (Time Expression Recognition and Normalization) 2004 competition, and the rule-based system HeidelTime (Strötgen/Gertz 2010) submitted by the TempEval-2 task (part of the SemEval 2010 workshop) which showed the best overall performance. The systems proposed for event annotation, in turn, include Klavans and Chodorow (1992), Siegel and McKeown (2001), Saurí et al. (2005), and Filatova and Hatzivassiloglou (2003), only to name a few. Finally, the systems for identifying temporal relationships are developed and described in Mani et al. (2006), Lapata and Lascarides (2006), and Chambers et al. (2007).

The existing datasets annotated with temporal information are represented by a number of corpora, including the TERN corpus[66] as well as the TimeBank corpus,[67] the AQUAINT

[65] An elaborated overview of the existing systems for temporal processing is provided in Marşic (2011)
[66] https://catalog.ldc.upenn.edu/LDC2005T07
[67] https://catalog.ldc.upenn.edu/LDC2006T08

corpus,[68] and the TempEval and TempEval 2 corpora.[69] The annotation (encoding) schema for temporal information developed so far are TIMEX and TIDES TIMEX2 (Ferro et al. 2000, 2005), STAG (Setzer 2001) as well as TimeML and ISO-TimeML (Pustejovsky et al. 2003).

6.2.3.5 Measuring Semantic Textual Similarity

The importance of finding related or similar sentences has been recognized, for instance, for the task of textual entailment recognition, including CD (Dagan et al. 2006) as well as question answering (Lin/Pantel 2001), paraphrase recognition (Dolan et al. 2004), and automatic essay grading (Attali/Burstein 2006). The task of recognizing similar or related sentences is termed as Semantic Textual Similarity (STS) in NLP.

A number of approaches for the STS task have been proposed up to now, relying on different methodologies such as, e.g., computing the string overlap (Baeza-Yates/Ribeiro-Neto 2011; Lin 2004) which is computed by means of similarity measures such as cosine similarity, Dice's coefficient, Euclidean distance, and Jaccard similarity, only to name a few. Though easy to realize, the string-based (or term-based) approach fails to recognize the similarity between sentences when these contain synonyms. Another approach measures the semantic similarity between texts by identifying the degree of similarity between their words, using lexical resources such as, e.g., WordNet. This approach is referred to as knowledge-based (Mihalcea, Corley/Strapparava 2006). Finally, another group of techniques computes the similarity between texts, computing the lexical similarity obtained from large corpora. These techniques are referred to as corpus-based (Mihalcea, Corley/Strapparava 2006). Together with knowledge-based measures, corpus-based measures represent the semantic similarity approach. Together, the preferred corpus-based similarity measures include the hyperspace analogue to language, or HAL (Lund/Burgess 1996), the latent semantic analysis, LSA (Deerwester et al. 1990; Landauer/Dumais 2008), the generalized latent semantic analysis, or GLSA (Matveeva 2008) as well as the explicit semantic analysis, ESA (Gabrilovich/Markovitch 2006). The latest studies (e.g. Kentner/de Rijke 2015) showed the good performance of distributed representation models (see Section 6.2.4), such as word2vec (Mikolov et al. 2013), for finding related sentences in a text corpus. An elaborated list of methods is provided in Gomaa and Fahmy (2013).

A unified framework for development and evaluation of the STS methods was created in a series of SemEval workshops on the STS task in the framework of SIGLEX workshop which is defined annually since 2012. The submitted methods and systems are summarized and

[68] https://catalog.ldc.upenn.edu/LDC2002T31
[69] http://www.timeml.org/timebank/timebank.html

compared in Agirre et al. (2012, 2013, 2014, 2015, 2016). The evaluation data on the wiki performance of the STS task can be freely obtained.[70]

The top systems from the SemEval 2012, 2014, and 2015 include an open source framework DKPro Similarity[71] (Bär, Zesch/Gurevych 2013) written in Java, TakeLab (Šarić et al. 2012) written in Python, and DLS@CU (Sultan, Bethard/Sumner 2015) written in Python.

6.2.4 Approaches to Meaning Representation

It is obvious that in order for the meaning of a linguistic expression to be processed by a machine, it must be captured in a machine-adaptable form, which in NLP, is referred to as meaning representation.

Up to now, a number of various approaches to meaning representation have been proposed. In the 1970s, the representation of linguistic meaning was under the influence of Montague's formal semantics (1974), which focused on the logical properties of natural languages. The intention of formal semantics was to define the rules which allow the transformation of a surface sentence form in the corresponding logical representation in a compositional way, concentrating on the possibility to represent logical features of natural language expression such as quantification, logical connectors, and modality. This logical representation is referred to as logical form (for more elaboration on logical form, see, e.g., Allen 1995; Jurafsky/Martin 2008; Schubert 2015). For instance, the logical form of the sentence *I have a car* constructed by means of the First Order Predicate Calculus (also called predicate logic) can be represented as $\exists x,y Having(x) \wedge Haver(S, x) \wedge HadThing(y, x) \wedge Car(y)$ (an example taken from Jurafsky/Martin 2000: 503). Further, the discourse semantics extended Montague's approach in order to apply it to the texts. A number of theories have been proposed, including, e.g., the Discourse Representation Theory, or DRT (Kamp 1981; Heim 1982) and the Segmented Discourse Representation Theory, or SDRT (Asher 1993; Lascarides/Asher 1993; Asher/Lascarides 2003). The representation of meaning by means of logic was widely applied for natural language processing, but due to its limitations such as the inability to deal with the lexical ambiguity, there was still a need for other more efficient possibilities of meaning representation.

An attempt to cope with the limitations of formal semantics with respect to meaning representation was made in lexical semantics which was proposed to represent the sentence surface by means of thematic roles such as agent, theme, recipient, location, etc., with respect to the argument structure of the verb (e.g. Fillmore 1968; Jackendoff 1987; Dowty

[70] http://ixa2.si.ehu.es/stswiki/index.php/Main_Page
[71] https://dkpro.github.io/dkpro-similarity/

1991). For instance, the sentence *John kicked the ball to the fence* can be represented as *∃e(kick(e) ∧ before(e, Now1) ∧ agent(e, John) ∧ theme(e, Ball2) ∧ goal-loc(e, Fence3))*, where e represents the kicking event (Schubert 2015). Such representation is also called neo-Davidsonian, following the view of Davidson (1967) that "verbs tacitly introduce existentially quantified events" (Schubert 2015). The difficulty of the approach appears as the result of the ambiguity in assigning of thematic roles (Riemer 2010).

Besides the logically motivated approaches to meaning representation have been developed, which tended to describe linguistic meaning by relating it to logic, psychologically inspired approaches, describing meaning from the perspective of human cognition modeling. Quillian (1968), for instance, proposed representing meaning based on a semantic network, a model for human associative memory, which has been successfully applied in a number of NLU systems (see Sowa 1987 for an overview). According to Quillian's approach, meaning can be constructed as a set of nodes representing concepts linked in a graph and graph edges which represent the relation between the concepts. Hereby, networks were used both for sentence meaning and knowledge representation. Moreover, Quillian developed simple operations to enable inferencing by means of a semantic network.

Further, inspired by the Quillian's semantic network and dependency theory, Hays (1964) and Schank (1972) developed the conceptual dependency theory which, in contrast to semantic network theory, aimed at describing content rather than structure. For this purpose, Schank (1972) defined a set of conceptual dependency primitives to describe actions – primitive acts (e.g. ATRANS refers to the transfer of ownership, possession, or control of an object), primitive conceptual categories (e.g. ACT refers to action), conceptual tenses (e.g. recipient (R): the receiver of an object; object (O): a thing that is acted upon), as well as diagrammatic conventions (arrows indicate the direction of dependency; double arrow indicates two way link between actor and action).

Figure 7 illustrates a conceptual dependency graph for the sentence *John pushed the cart*. The conceptual dependency graphs have been used in the NLU systems described in Schank (1975) and Schank and Abelson (1977).

Figure 7: Meaning of *John pushed the cart* as a conceptual dependency graph.[72]

[72] PROPEL: application of physical force to an object.

Sometime later, Minsky (1975) proposed representing knowledge (also applicable to the representation of sentence meaning) based on so-called frames. In linguistics, the research on frames, both from theoretical and methodological view, was influenced by Fillmore (1968). The main idea of the frame representation theory as formulated in Minsky (1975: 212) is as follows:

> When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a structure called a Frame. This is a remembered framework to be adapted to fit reality by changing details as necessary. A frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed.

For example, the possible frame for the concept of *Pacific island* can be construed as follows: [is-a: island, located: Pacific_ocean, belongs_to: country, name: island_name], where is-a, located, belongs_to and name are the slots in the frame which represent the concept (the example is taken from Ovchinnikova 2012: 29). Further, building on this idea, Schank and Abelson (1977) developed the notions of scripts, plans, and themes to deal with the procedure (Section 3.4.3.1).

Along with the growing popularity of machine learning techniques in NLP, statistical approaches to meaning representation have been developed. One of these approaches is based on the idea expressed by Firth (1957) stating, “you shall know a word by the company it keeps”, similar to the quotation of Wittgenstein (*Philosophical Investigations*, 43 (1953)) “the meaning of a word is its use.“ In NLP, this idea is also referred to as the distributional hypothesis as it goes from the assumption that there is a correlation between similarity of meaning and similarity of distribution or “words which are similar in meaning occur in similar contexts” (Rubenstein/Goodenough 1965: 627). The notion of the distributional hypothesis is considered to have originated in Harris (1954). “If we consider oculist and eye-doctor we find that, as our corpus of actually occurring utterances grows, these two occur in almost the same environments (….). If A and B have almost identical environments except chiefly for sentences which contain both, we say they are synonyms: oculist and eye-doctor” (Harris 1954: 156-157).

Up to now, a number of distributional semantic models (DSM, also word space models, or distributional similarity models) have been proposed. A DSM can be defined as a scaled and/or transformed co-occurrence of matrix *M* with rows representing the target terms, and their distribution in contexts and collocates is listed in columns. The target terms can be a word form, lemma, phrase, morpheme, word pair, etc., such as illustrated in Table 28.

	get	See	Use	Hear	eat	Kill
knife	0.027	-0.024	0.206	-0.022	-0.044	-0.042
cat	0.031	0.143	-0.243	-0.015	-0.009	0.131
dog	**-0.026**	**0.021**	**-0.212**	**0.064**	**0.013**	**0.014**
boat	-0.022	0.009	-0.044	-0.040	-0.074	-0.042
cup	-0.014	-0.173	-0.249	-0.099	-0.119	-0.042
pig	-0.069	0.094	-0.158	0.000	0.094	0.265
banana	0.047	-0.139	-0.104	-0.022	0.267	-0.042

Table 28: A DSM for the concept *dog*. *Note*: From DSM Tutorial (Stefen Evert et al. 2009-2016)[73].

The widely-applied DSM include LSA, HAL as well as different variants of topic models (e.g., the topic-based vector space model described in Becker/Kuropka 2003; the LDA developed by Blei et al. (2003); the generalized vector space model proposed in Tsatsaronis/Panagiotopoulou 2009, to name only a few).

At present, a popular approach to meaning representation is a distributed representation (also called word embeddings) of a meaning, which is not to be confused with distributional representation. Though both are built on the distributional hypothesis, distributed representation is a low-dimensional vector representation which can be obtained from models based on the neural network (e.g., the Collobert and Weston embeddings described in Collobert and Weston (2008), the HLBL embeddings (Mnih/Hinton 2009), and word2vec (Mikolov et al. 2013)) and the matrix factorization-based model such as, e.g. the Stanford Glove model (Pennington, Socher/Manning 2014). A concept with respect to distributed representation is represented by many vectors. Distributional representation, in turn, is the distribution of words in their specific context, obtained from the word-context matrix. In contrast to distributional approach, distributed representations, due to their complexity, are more difficult to construct, but they are not as memory intensive.

6.2.5 Computational Sources of Knowledge

6.2.5.1 Lexical Resources

In Section 6.2.2.1, we have already addressed the lexical resources VerbNet, PropBank, FrameNet, NomBank, which are the essential databases for the development of methods for an SRL task.

WordNet is considered to be the most popular lexical resource at the present time. Originally developed for English by Miller (Miller et al. 1990) and described in Fellbaum (1998), the WordNet exists at present for a number of languages, including German, Russian, Chinese,

[73] http://esslli2016.unibz.it/wp-content/uploads/2015/10/dsm_tutorial_part1.slides.pdf

Japanese, among others (for a complete list of supported languages, visit the website of the Global WordNet Association).[74]

In general, the WordNet can be compared with a thesaurus grouping the words according to their meanings. The lexical-semantic knowledge in WordNet is organized as a network. Nouns, verbs, adjectives, and adverbs are grouped into units, so-called cognitive synonyms, or synsets, representing a distinct concept, including all its senses, definitions, and short examples. For example, the multiple senses of the noun *dog* are represented by seven synsets as illustrated in Figure 8.

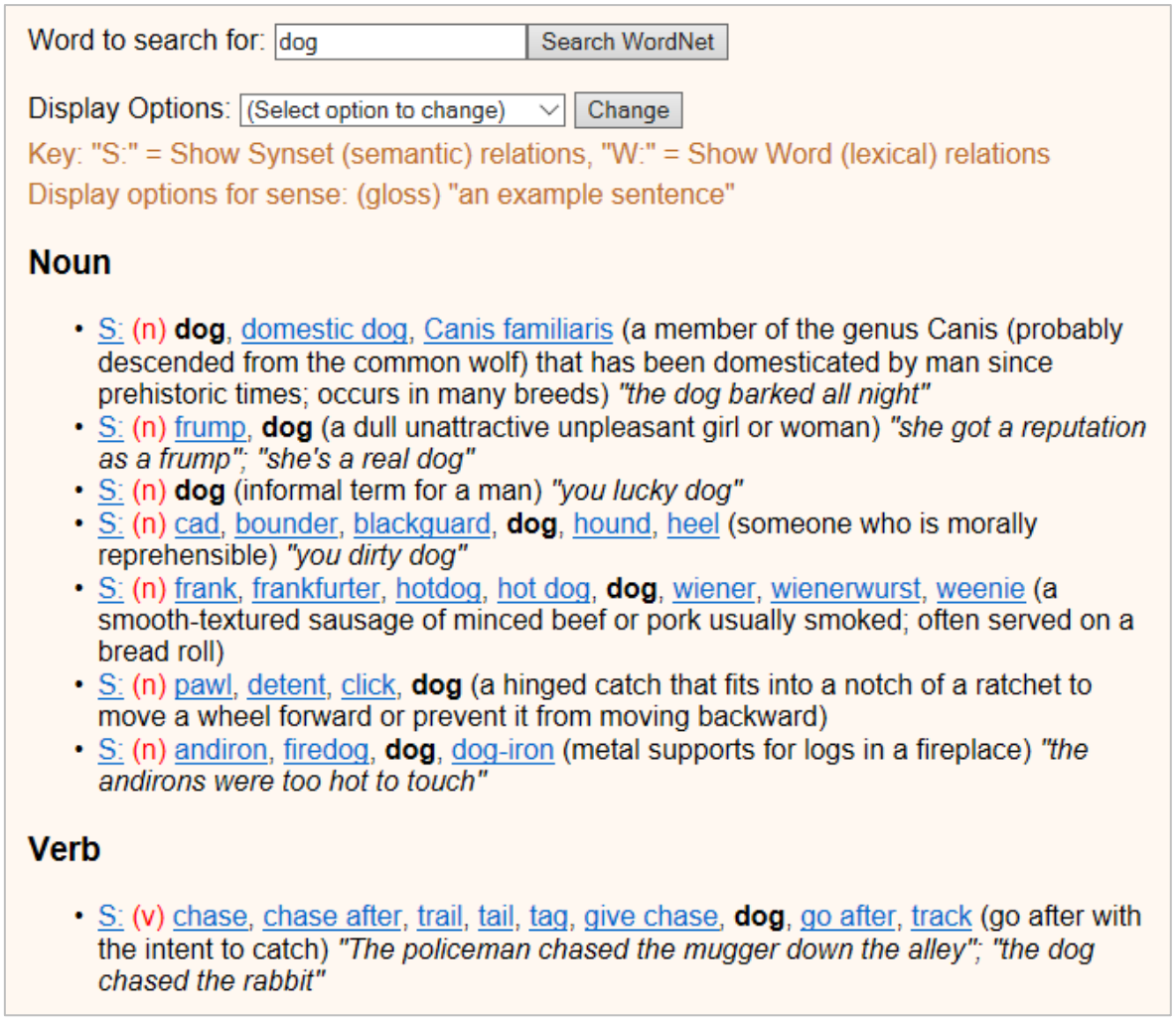

Figure 8: The multiple senses of the noun *dog* as represented in WordNet.

The synsets are further interconnected with each other by means of lexical and conceptual-semantic relations. It is to be noted that only the synsets of the same parts-of-speech are linked to each other in WordNet, dividing the resource into four parts: WordNet for nouns, WordNet for verbs, WordNet for adjectives, and WordNet for adverbs. The noun synsets,

[74] http://globalwordnet.org/wordnets-in-the-world/

for instance, are interconnected by means of a super-subordinate relation (also called hyperonymy, hyponymy, or is-a relation, e.g., *canine* is a hypernym of *dog* and *dog* is a hyponym of *canine*), the part-whole relation (holonymy and meronymy, e.g. *engine* is a meronym of *car*), and term coordination (e.g., *wolf* is a coordinate term of *dog* and *dog* is a coordinate term of *wolf*). Adjective synsets are interconnected by means of antonymy, etc. Verbs are organized in WordNet according to the troponym relation (or according to the specific manner of characterizing an event, including volume such as in *{communicate}-{talk}-{whisper}*, speed *{move}-{jog}-{run}* or intensity of emotion *{like}-{love}-{idolize}*), hypernym (or a kind of relation, e.g. *to perceive* is a hypernym of *to listen*), entailment (e.g. *to snore* entails *to sleep*), and term coordination (verbs share a common hypernym, e.g. *to lisp* and *to yell*). The only semantic relation which connects the synsets of different parts of speech is the attribute-value relation which in the English WordNet 3.1 version has been established between the verb-noun pairs based on the morphological similarity of the words such as in the case of observe (verb), observant (adjective), and observation, observatory (nouns).

Though WordNet is widely used for different NLP applications, there are some critique aspects towards the construction of the resource. Agirre and Lopez de Lacalle (2003), for instance, criticize the fine-granularity of the word senses which, in turn, complicates the process of sense distinction. Oltramari et al. (2002), in turn, address the conceptual inconsistency in WordNet, which can lead to incorrect reasoning, and they propose a methodology for dealing with the hypero-hyponymical relations.

The current WordNet English as a full package, including WordNet browser, the command-line tool, and the database files with the InstallShield self-extracting installer for Windows is available as version 2.1; for UNIX, Linux, Mac OS X, and Solaris, it is available as version 3.0; and for the database files only, it is version 3.1. The statistics on English WordNet 2.1, 3.0, and 3.1 are summarized in Table 29. The source is freely available, it can be navigated in a browser,[75] can be downloaded,[76] or can be accessed within the NLTK package.

	WordNet 2.1			**WordNet 3.0**			**WordNet 3.1**		
POS	Total no. of words	Synsets	No.word-sense pairs	Total no. of words	Synsets	No.word-sense pairs	Total no. of words	Synsets	No.word-sense pairs
Noun	117,097	81,426	145,104	117,798	82,115	146,312	n.a.	n.a.	n.a.
Verb	11,488	13,650	24,890	11,529	13,767	25,047	n.a.	n.a.	n.a.
Adjective	22,141	18,877	31,302	21,479	18,156	30,002	n.a.	n.a.	n.a.
Adverb	4,601	3,644	5,720	4,481	3,621	5,580	n.a.	n.a.	n.a.
Total	155,327	117,597	207,016	155,287	117,659	206,941	n.a.	n.a.	n.a.

Table 29: WordNet 2.1, 3.0 and 3.1 database statistics.

[75] http://wordnetweb.princeton.edu/perl/webwn
[76] http://wordnet.princeton.edu/wordnet/download/

DIRT (Discovery of Inference Rules from Text) refers both to an algorithm and to a collection of paraphrases which have been automatically learned from 1 Gb of newspaper text corpora, including the San Jose Mercury, the Wall Street Journal, and AP Newswire from the TREC-9 collection. The main idea underlying the algorithm is that it automatically learns paraphrases from texts, following the distributional hypothesis (Section 6.2.4) over paths in dependency trees. The algorithm and resource have been developed by Lin and Pantel at the University of Alberta. More details on the resource and algorithm are provided in Lin and Pantel (2001). The current version of the DIRT consists of around 231,000 unique patterns. The top 20 paraphrases of *x solves y* include: *Y is solved by X*, *X resolves Y*, *X finds a solution to Y*, *X tries to solve Y*, *X deals with Y*, *Y is resolved by X*, *X addresses Y*, *X seeks a solution to Y*, *X does something about Y*, *X solution to Y*, *Y is resolved in X*, *Y is solved through X*, *X rectifies Y*, *X copes with Y*, *X overcomes Y*, *X eases Y*, *X tackles Y*, *X alleviates Y*, *X corrects Y*, *X is a solution to Y*, *X makes Y worse*, and *X irons out Y*. The DIRT resource is freely available for research purposes and can be obtained from its developers.

Another resource, developed by Patrick Pantel in a team with Chklovski, is VerbOcean (Chklovski/Pantel 2004). VerbOcean is a broad-coverage semantic network of verbs. Currently, the resource consists of 3,477 unique verbs with 22,306 relations established between the verbs. The types of VerbOcean semantic relations include similarity with 11,515 instances (e.g. *produce – create*), strength with 4,220 instances (e.g. *wound – kill*), antonymy with 1,973 instances (e.g. *open – close*), enablement with 393 instances (e.g. *fight – win*) and happens-before with 4,205 instances (e.g. *marry – divorce*, *buy – own*). The relations are additionally provided with information to whether they are transitive and symmetric. VerbOcean, of which the refined version is currently not available, can be freely downloaded from the website of the project.[77]

WikiRules! (Shnarch, Barak/Dagan 2009) is a database which contains about 8 million of lexical reference rules extracted from Wikipedia. Hereby, a lexical reference rule is defined as a directional relation which identifies a concrete reference from its left-hand-side (LHS) to its right-hand-side (RHS) such as in the cases of *Margaret Thatcher* → *United Kingdom*, A*bbey Road* → *The Beatles*, etc., indicating more general relations than the traditional meaning relations (synonymy, hyponymy etc.). It is to be noted that LHS represents a title of the Wikipedia article, while RHS is a definition term extracted from the first sentence of the article. For this reason, the rules of the WikiRules! include mainly named entities and terminological terms, which is a characteristic of encyclopedias. According to the

[77] http://demo.patrickpantel.com/demos/verbocean/

developers, the database shows comparable performance to the WordNet lexical database. The resource is freely available at the website of the Natural Language Processing Lab of the Bar-Ilan University.[78] It can be imported into the MySQL DB database or downloaded as .csv files.

6.2.5.2 Ontologies

The notion of *ontology* initially originates from metaphysics and, in a philosophical sense, refers to the study of the nature of being, existence, and reality. Since the 1970s, the term also refers to the computational models of knowledge developed in the field of Artificial Intelligence (AI). The most accepted definition of ontology in AI, proposed in Gruber (1993: 199), regards the latter as "an explicit specification of a conceptualization". Further, Guarino (1998: 5) builds his definition on this view, proposing a formal account for the notion of ontology. He defines ontology as

> "(…) a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world. The intended models of a logical language using such as vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models"

referring to a set of axioms with logical theory and with language to an ontology representation language such as RDF (Resource Description Framework) and OWL (Web Ontology Language), rather than natural language. According to Heflin and Pan (2004: 63), "an ontology provides a common vocabulary to support the sharing and reuse of knowledge".

In the following, the ontologies YAGO, DBpedia, and FreeBase will be presented, containing information about named entities and created from structural sources such as Wikipedia as well as a ConceptNet (Speer/Havasi 2012) ontology generated from existing knowledge resources. Other ontologies which will not be further addressed here include DOLCE (The Descriptive Ontology for Linguistic and Cognitive Engineering; Masolo et al. 2003), SUMO (The Suggested Upper Merged Ontology), and OpenCyc (e.g. Lenat/Guha 1990).

Wikipedia is an open-source encyclopedia which is collaboratively being created and edited by volunteers. The current version of Wikipedia for English contains 5,544,517 articles and it averages 600 new articles per day. Versley, Poesio, and Ponzetto (2016: 400) name at least three reasons that explain the attractiveness of Wikipedia as a knowledge repository. These are as follows:

- A good coverage across many domains: it contains a large amount of information, in particular at the instance level;

[78] http://u.cs.biu.ac.il/~nlp/resources/downloads/lexical-reference-rules-from-wikipedia/

- Multilingual: it is available with a (mostly) uniform structure for hundreds of languages, even though the size of Wikipedia in different languages vary substantially (…);
- Up-to-date: it includes continuously updated content, which provides current information.

Since 2004, the Wikipedia articles on similar subjects are grouped together by means of assigning categories to the article under [[Category:XYZ]]. Each article can be assigned to one or more categories, and the categories, in turn, can contain subcategories. The categories can be found at the bottom of each article. By clicking on them, similar articles will be found listed. Beside a category assignment, Wikipedia articles also contain structural information in form of infoboxes which are built on a common template, tables, images depicting the topic of the article, links to external web pages, and to other Wikipedia articles. In order to use Wikipedia for natural language processing, that is, to extract structural information, the need has been recognized of creating such knowledge databases as DBpedia (Mendes et al. 2011) and YAGO (Suchanek, Kasneci/Weikum 2007).

DBpedia is a joint project of the University of Leipzig University, the University of Mannheim, and OpenLink Software which enables processing of queries such as, for instance, *Give me all cities in New Jersey with more than 10,000 inhabitants* or *Give me all Italian musicians from the 18th century* on basis of Wikipedia content. It uses the SPARQL query language to query this data. The English version of DBpedia at present describes 4.58 million instances extracted from Wikipedia, and hereby, 4.22 million instances are classified in an ontology including 1,445,000 persons, 735,000 places (with 478,000 populated places), 411,000 creative works (e.g., 123,000 music albums, 87,000 films, and 19,000 video games), 241,000 organizations (e.g., 58,000 companies and 49,000 educational institutions), 251,000 species, and 6,000 diseases. DBpedia uses the RDF to represent the information extracted. Figure 9 Illustrates an excerpt from the DBpedia ontology. For example, William Shakespeare can be an instance of this representation (Writer → Artist → Person) as well as his work *Much ado about nothing* (Book → Work), which are considered to be related instances via property, author.

DBpedia is a free source licensed under Creative Commons Attribution-ShareAlike 3.0 License and the GNU Free Documentation License and can be downloaded from the website of the DBpedia project.[79] The latest DBpedia release is DBpedia version 2016-10.

[79] http://wiki.dbpedia.org/

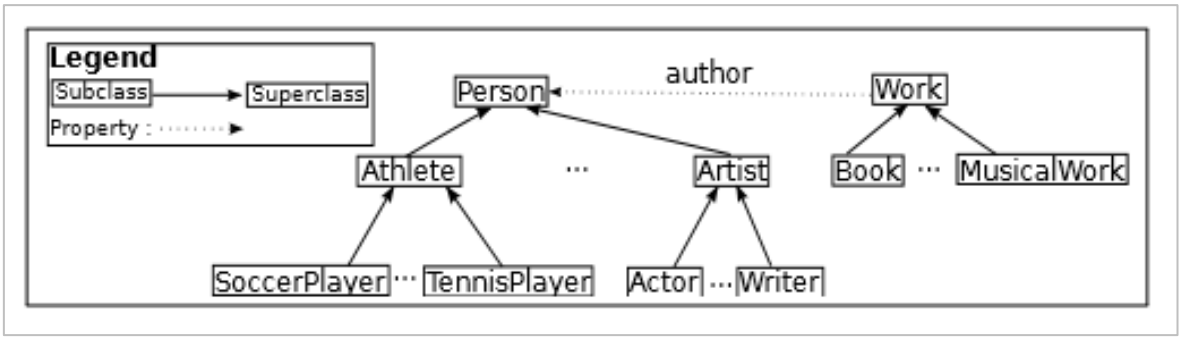

Figure 9: Excerpt of the DBpedia. *Note*: From DBpedia and the live extraction of structured data from Wikipedia (Morsey et al. 2012: 5).

Another source for utilizing Wikipedia's articles is YAGO (Yet Another Great Ontology), developed in a joint project of the Max Planck Institute for Informatics and the Telecom ParisTech University, and described in Suchanek et al. (2007; YAGO1), Hoffart et al. (2011; YAGO2), Hoffart et al. (2013; YAGO2), Biega, Kuzey, and Suchanek (2013; YAGO2), and Mahdisoltani, Biega, and Suchanek (2015; YAGO3). It additionally incorporates WordNet and the geographical database GeoNames.[80] The current version of the YAGO database contains more than 10 million named entities, such as persons, organizations, and countries, and more than 120 million facts about these entities. One of the advantages of the ontology is that it attaches a temporal and a spatial dimension to many of its facts and entities. The ontology has been manually evaluated and showed a 95% accuracy. YAGO is open-source and can be freely downloaded or explored online.[81] Also, the source code of YAGO is freely available.

ConceptNet[82] is an associative network of knowledge on the meanings of words, developed within the Open Mind Common Sense Project initiated in 1999 at the MIT Media Lab. For instance, the knowledge on the concept of the type event *eat sandwich* contained in ConceptNet includes the information illustrated in Figure 10. The current version of ConceptNet 5.5 contains 1.6 million assertions, 300,000 concepts of four types, including events (e.g., *eat sandwich*), things (e.g., *morning coffee*), places (e.g., *near school*), properties (e.g., *dark*, *very expensive*), 20 binary relations (e.g., *CapableOf* as in *CapabaleOf: dentist*, *pull tooth*), and incorporates knowledge from other sources such as DBpedia, OpenCyC, Wiktionary, WordNet as well as a word game, Verbosity.

[80] http://www.geonames.org/
[81] https://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/yago/
[82] http://conceptnet.io/

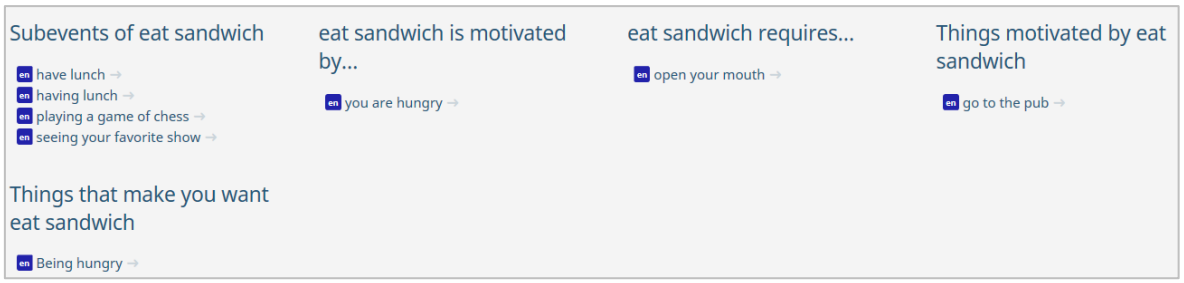

Figure 10: The knowledge of the concept *eat sandwich* as represented in ConceptNet.

The sources can be applied for multiple languages, including English, German, Russian, Chines etc. It is a free source, and it can be freely downloaded[83] or used with the Web API. ConceptNet 5 is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Portions of the source are also available under a Creative Commons Attribution 4.0 International License.

6.3 Summary

The topic of Chapter 6 was the conceptual design of a CD system by considering the previously elaborated theory, as well as presenting the natural language processing tasks which are involved in a general process of meaning processing and interpretation at lexical, morphological, syntactic, semantic, pragmatic, and discourse levels and can contribute to an effective automatic detection of contradictions. For each task, the main approaches and the state-of-the-art tools have been presented.

Thus, the tasks addressed at lexical, morphological, and syntactic levels include tokenization and sentence segmentation, which are prior to a machine-based processing, stop words detection, and removing for the elimination of irrelevant noisy words with less or no content-contributing. By this means, it contributes toward more meaningful results of part-of-speech tagging which is an automatic assigning of parts of speech to the words, stemming and lemmatization which are the tasks for the transformation of the words to their initial form as well as parsing and chunking which refer to an identification of syntactic relations of constituents in a sentence.

At the semantic, pragmatic, and discourse levels, the presented tasks were semantic role labeling which is a task for an identification of thematic roles, the recognition of textual entailment as well as anaphora/coreference resolution.

Further tasks addressed were named entity recognition, negation, modality, and time processing, sentiment analysis which is the process of identification of positive, negative, and

[83] https://github.com/commonsense/conceptnet5/wiki/Downloads

neutral opinions, attitudes, and sentiments as well as measuring the semantic textual similarity which is a task of finding related sentences, texts, and text passages.

Besides these tasks, the main approaches to meaning representation discussed above include the logical form, the meaning representation by means of thematic roles, semantic network, conceptual dependency graph and frames as well as distributed and distributional models. Finally, computational sources of knowledge such as WordNet, VerbOcean, VerbNet, PropBank, NomBank, FrameNet, DIRT, WikiRules!, and a number of ontologies have been presented.

7 Physical Design of a CD System and Implementation

The present chapter incorporates the theory introduced in the previous chapters and based on it, proposes a system named Contradictio for detecting the contradictions and contrarieties occurring in texts of online news articles written in the English language. The chapter begins first with a sketch of the architecture of a prototype CD system and describes its main components, or modules and outlines the main tasks which are included in the modules (Section 7.1). The idea of contradiction detection underlying the Contradictio system can be described as a process of finding in one or more news texts at least two declarative sentences or parts of a declarative sentence which potentially have the same semantic content and refer to the same situation in the world in the same respect as well as include elements which are contradictory or in contrary relation to each other. The need to introduce and operate with the concept of potential contradiction, rather than with that of actual contradiction, is discussed in Section 7.1 as well. The implementation of the Contradictio system is described in Section 7.2. The dataset used by the implementation and further by the evaluation of the system has been previously introduced in Section 5.3. Finally, the evaluation of the system is the topic of Section 7.3.

7.1 System Architecture and Potential Contradiction

Considering the conceptual design as addressed in the previous chapter, we propose a CD system which consists of four modules performing preprocessing, finding two parts which may potentially constitute a contradiction, filtering of non-relevant sentence pairs, and making the final decision on the presence of contradiction/contrariety. The setup, or architecture of the system is illustrated in Figure 11.

The main purpose of the module *Preprocessing* is to collect linguistic information and prepare the textual data for processing in the subsequent modules. The collected information on the data is saved as a knowledge graph, storing among others information on grammatical number, thematic role, part-of-speech, original form of the token and its lemma as well as a normalized form, tense, and aspect in case of verbs as well as co-referent expressions in the sense of computational linguistics (without knowledge whether the expressions refer to the same object in the real world) for each token. That is, not only the tasks of tokenization, sentence splitting, lemmatization, and part-of-speech tagging but also stop words removing and semantic role labeling are fundamental here. Moreover, the step of normalization in order to transform numerical information on time, date, quantity, and quality to a common base is conducted in this module. The identification of the referents is the task of the next module.

Following the step of the collection of information on the textual data, the main purpose of the module *Finding parts of contradiction* is to detect in a text or a corpus of texts a pair of thematically related sentences or parts of a sentence which may contain a contradiction/contrariety cue and check whether the found sentences fulfill the first of the necessary conditions of contradiction (see Section 3.6).

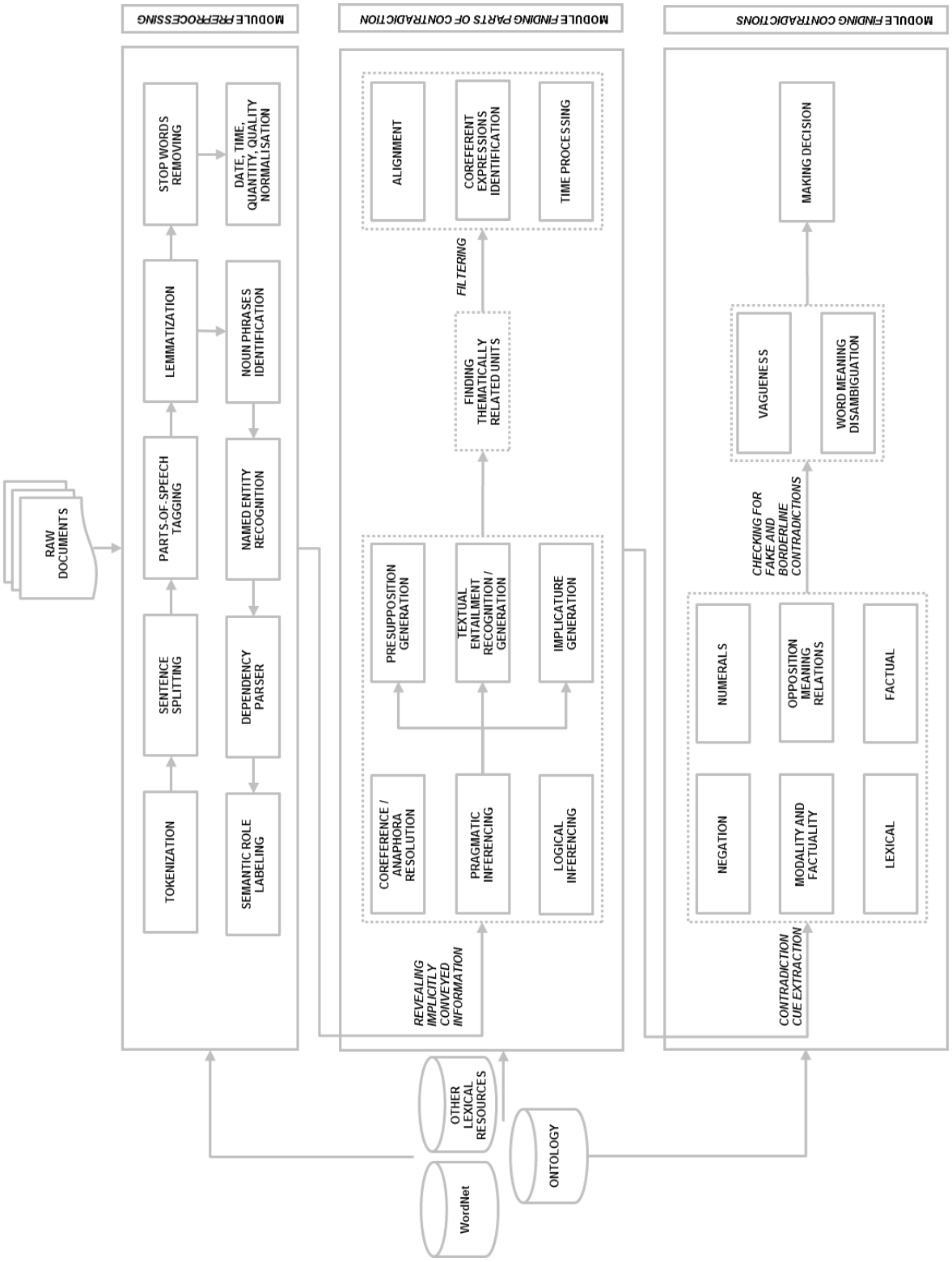

Figure 11: The architecture of the Contradictio system.

The module *Finding parts of contradiction* consists of three submodules including *Revealing implicitly conveyed information*, *Finding thematically related units* (in our case, sentences), and *Filtering*.

The main purpose of the submodule *Revealing implicitly conveyed information* is to reveal information which is implicitly conveyed in the text/texts. That is, the steps to conduct here

are the anaphora/coreference resolution, pragmatic inferencing including recognition/generation of presupposition, entailment, and implicature as well as the step of logical inferencing. Whether the overall performance of the system can profit from the task of anaphora/coreference resolution or whether this task can be left out will be examined in Section 7.2. Thematically related sentences are to be identified in the submodule *Finding thematically related units*. These are then aligned and further analyzed in the submodule *Filtering* in order to filter out those sentence pairs which do not satisfy some of the necessary contradiction conditions. That is, the main tasks conducted in this submodule include first a determining of whether the events described in the sentences have the same time reference. Second, the decision is to be made on whether the described events are co-referent, that is, refer to the same object in the real world. In the present study, we follow a heuristic approach to fulfill this task by assuming that two expressions are co-referent when the lexical chains they occur in are similar.

Finally, in the module *Finding contradictions*, the decision is to be made on whether the sentences/or parts of a sentence found are in contradictory/contrary relation with each other. For this purpose, the system searches for the contradiction/contrariety cues identified in Chapter 5 by processing the modality and factuality of the sentences, identifying the negation and opposition meaning relations, finding divergent numerical information, slightly different spelling of the same names, and use of incorrect encyclopedic knowledge, as well as by identifying the lexical meaning. Finally, the analyzed sentences are classified into the categories *yes* for a (potential) contradiction and *no* in case of a non-contradiction.

When taking into consideration the complexity of natural language contradictions, including the previously discussed phenomena of vagueness and lexical/contextual ambiguity which are able to cancel a contradictory/contrary relation or make a contradiction to a matter of interpretation, it appears to be ambitious and to some degree reckless to speak about automatic detection of *actual* rather than *potential* contradiction. For this reason, we propose to introduce the category *potential contradiction* when addressing the *yes*-class to cover borderline and fake cases and to cope with gaps including truth value and reference fixation.

We define the term potentiality, following the theoretical explication of the dichotomy of *actuality-potentiality* discussed by Aristotle in *Metaphysics*. Following Aristotle, *actual* contradiction refers to "anything which is currently happening", while *potential* contradiction refers to something that “might chance to happen or not to happen” (*Metaphysics* 1019a - 1019b).

Besides the phenomena of vagueness (borderline contradiction) and lexical/contextual ambiguity (fake contradiction), also the fixation on the references and as a result, a decision on the coreference of the expressions can be regarded as a limitation of natural language processing at the present time. For this reason, the proposed implementation of the system

(Section 7.2) reveals to be based on a reduced view of a complex nature of contradiction being unable to cope with a full range of aspects, which can be decisive for the correct system's decision on contradiction. That is, the decision of the system is based on the defined conditions of the contradictions that are necessary for a system to predict a possible, or *potential* contradiction but *not enough* to state that the sentences constitute an *actual* contradiction. Here, an additional human judgment is required – the confirmation or verification of the (actual) contradiction status of a pair of sentences which are predicted by the system as potential contradictions.

7.2 System Implementation

7.2.1 Module Preprocessing

The implementation of the module *Preprocessing* was initiated by the uploading the input data which are the raw English news texts. In total, 165 news texts contained in the corpus were processed. The present module, as well as subsequent modules, were realized by means of the programming language Python and by the application of existing lexical resources and NLP tools. These were accessed within existing Python libraries for natural language processing (e.g. NLTK) or by correspondent Python APIs if that was required.

After the data was uploaded, the texts were tokenized and based on the results of the tokenization, the sentences were split. In order to evaluate the results of the system's application, in the next section, each sentence in the corpus was initially provided with information as to the text it occurs in (text ID) and its numerical order in the text (sentence ID).

Further, to reduce the inflectional forms of each string token into a common base, the process of lemmatization was applied. As the inflectional forms such as e.g. the plural form can be significant for the detection of contradictions, both the lemmatized and original unlemmatized data were also saved. For realizing the tokenization, sentence splitting, and lemmatization, the tokenization function from the Python NLTK library as well as the TreeTagger tool (s. Section 6.2.1.3) for English were applied. TreeTagger was accessed by a Python wrapper, treetaggerwrapper 2.2.4,[84] which is provided by Laurent Pointal and released under the General Public License. The wrapper can be used independent of the operating system and can be used under both Python 2 and Python 3.

Besides the above-described tasks, the TreeTagger tool was additionally applied in the next step to tag the data with their parts-of-speech. This was done in order to extract the words with high semantic content (noun, verb, adjective), by this means eliminating the stop words,

[84] https://pypi.org/project/treetaggerwrapper/

such as the articles *a* and *the*, which have little or no semantic content. In contrast to a common approach of removing stop words based on pre-defined lists, we followed another approach, not by eliminating the stop words, but by extracting words with high semantic content. Moreover, besides stop words, removing the results of POS tagging was used for the purpose of finding open compound nouns such as *Los Angeles*, *United States of America*. For this task, chunking was conducted following our own pre-defined grammar rules.

Further, in order to collect information about the thematic roles of the constituents in the sentences, a semantic role labeling was conducted. For this purpose, an open source pythonic library, practNLPTools,[85] was used, which is based on SENNA and the Stanford Dependency Parser Extractor. SENNA is a software which provides functions to conduct a number of NLP tasks, including part-of-speech tagging, chunking, name entity recognition, semantic role labeling and syntactic parsing, among others. The software is distributed under a non-commercial license. Further, it should be noted that SENNA and the Stanford CoreNLP tools are considered to be the most effective NLP tools at present.

As news texts are characterized by a frequent occurrence of numbers and figures, the step of normalization was conducted. The information on time, date, quality, and quantity, especially expressing currency and measurement, was transformed into a common base. Moreover, numerals, primarily those occurring along with adverbs of degree, were presented as ranges as proposed in de Marneffe et al. (2008). According to this approach, the system should be able to recognize that concepts such as *more than five* and *six* are not contradictory. The adverbs of degree considered in the present study include *about*, *approximately*, *at least*, *more than*, *up to*, *almost* and *less than*.

All information collected in this module was stored in a graph as illustrated in Figure 12 for each token of the sentence. Thus, the graph for nouns includes the information on the parts of speech, noun class (proper noun or common noun), lemma, grammatical number (singular, plural), syntactic role, thematic role, co-referents, the presence of a determiner (the definite and indefinite articles *the* and *a* or *an*, demonstratives *this*, *these*, and *that*, possessive determiners e.g. *my*, *their*, quantifiers e.g. *many*, *few*, numerals, distributive determiners *each*, *any*, and an interrogative determiner *which*) as well as the presence of negation. Similar to the noun graph, the graph for verbs also includes information on the parts of speech, lemma, and the presence of negation but also information on aspect, tense, voice, mood, person and number, the type of the verb (auxiliary, modal), and the class of the Verb according to Levin's classification of verbs (Levin 1993). The graph for adjectives includes information on the lemma, form (simple, comparative, superlative) and the presence of

[85] https://github.com/biplab-iitb/practNLPTools/blob/master/practnlptools/tools.py

negation. In turn, the graph for pronouns contains the categories lemma, type (personal, possessive), co-referents, and presence of negation, the graph for adverbs include the categories class (postnominal adverbs, sentence adverbs (*however*, *but*), adverbs of manner, adverbs of frequency, adverbs of place and time, adverbs of degree, adverbs modifying the whole sentence), form (simple, comparative, superlative) and the presence of negation. Finally, the graph for numerals includes information on the type of numeral (cardinal numbers, numeral nouns, fractions, ordinal numbers) and its normalized form. An extension of the graphs, with information on co-referents, modality, negation, and factuality, will be made in the subsequent modules.

Sentence unit	**Obama**
Part-of-speech	Noun
Noun class	Proper
Lemma	Obama
Grammatical number	Singular
Syntactic role	Subject
Thematic role	Agent
Co-referents	Barack, Barack Obama, president, president of the USA, USA president
Determiner	Zero
Negation	No

Figure 12: An information graph for the noun Obama from the sentence, *Obama speaks to the media in Illinois*.

7.2.2 Module Finding Parts of Contradiction

Following the step of data preprocessing and based on the collected information, the subsequent module *Finding parts of contradiction* aims at finding thematically related sentences which may potentially contain a contradiction/contrariety cue as identified in Section 5.2. The module consists of three submodules. These are *Revealing implicitly conveyed information*, *Finding thematically related units* (in our case sentences), and *Filtering*.

First, the submodule *Revealing implicitly conveyed information* was realized. It is to be noted that only the anaphora/coreference resolution for the purpose of identifying the co-referents of the pronouns and noun phrases in the text was applied in this submodule. The tasks of pragmatic and logical inferencing contained in the submodule were left beyond the scope of system implementation at present by this means defining the tasks for future work.

As already mentioned above, the task of anaphora/coreference resolution is referred to in the sense of computational linguistics, meaning the identification of co-referents without knowing whether they co-refer in the real world. For performing the anaphora/coreference resolution, the CorefAnnotator of the Stanford CoreNLP library was applied. All Stanford CoreNLP tools are written in Java, and for this reason, the library was accessed by an open-

source Python wrapper to the Stanford server 3.4.1 stanford-corenlp-python,[86] developed by Dustin Smith. All Stanford tools applied in this and subsequent modules were run as a server since Stanford CoreNLP uses a number of trained models which require a large machine memory capacity. Whether the anaphora resolution has a positive impact on the overall performance of the system will be evaluated in the subsequent modules.

For implementing the submodule *Finding thematically related units*, we proposed applying the word mover's distance model (WMD), which is the model currently preferred for finding thematically similar sentences. The degree to which the model is suitable for the task of finding thematically related sentences potentially containing a contradiction/contrariety cue will be discussed in the next section.

According to the results reported in Kusner et al. (2015), the WMD model utilizing word2vec embeddings showed the best performance, with a low error rate in measuring textual similarity (in case of the WMD, it is correct to speak of textual dissimilarity measuring) in comparison to bag-of-words, tf-idf (Salton/Buckley 1988), Okapi BM25 (Robertson et al. 1995), LSI (Deerwester et al. 1990), LDA (Blei et al. 2003), mSDA (Chen et al. 2012), and the CCG (Perina et al. 2013) baselines underlying textual similarity computation with Euclidean distance for kNN classification.

In general, the choice of WMD can be explained by several reasons which are the properties of the model as defined in Kusner et al. (2015) on one side and model capabilities with regard to the task of textual similarity measuring on the other.

Thus, Kusner et al. (2015: 2) in particular emphasize the following properties of the model: "1. It is **hyper-parameter free**[87] and straight-forward to understand and use; 2. It is highly **interpretable** as the distance between two documents can be broken down and explained as the sparse distances between few individual words; 3. It naturally incorporates the knowledge encoded in the *word2vec* space and leads to **high retrieval accuracy** – it outperforms all 7 state-of-the-art alternative document distances in 6 of 8 real-world classification tasks."

Regarding the capabilities of the WMD for measuring textual similarity (and dissimilarity), the following aspects can be particularly emphasized. First, the algorithm does not require incorporating lexical resources in order to find related words and phrases. Second, the metric is suitable in particular for short texts (sentences). Third, the metric is able to find similar/related sentences which do not contain the same keywords. Thus, the algorithm can

[86] https://github.com/dasmith/stanford-corenlp-python
[87] The original font of the preceding text was preserved.

determine that *Obama speaks to the media in Illinois* and *The president greets the press in Chicago* convey the same information. However, the algorithm follows a heuristic approach as it cannot identify whether *Obama* in the first sentence and *the president* in the second are really co-referent expressions, that is, whether they really refer to the same person in the real world. As a disadvantage of the WMD, which uses vector representation of the texts as a baseline computed, for instance, with word2vec, the need of a large amount of textual data can be mentioned and, as a consequence, of a large machine memory capacity to learn word embeddings effectively and efficiently.

In general, the WMD was inspired by the Earth mover's distance, which is a metric developed for the purpose of finding similar images (Rubner et al. 1998). The main idea behind the WMD can be described as finding the distance between texts T_1 and T_2 as the cumulative sum of the minimum distance that each word vector of the word vector representation of text T_1 needs to travel in a vector space to the closest word vector of the word vector representation of text T_2 as illustrated in Figure 13. As reported in Kusner et al. (2015), the WMD utilizes a representation of texts as weighted embedded words. These can be computed using a word2vec algorithm by the skip-gram model proposed in Mikolov et al. (2013). Beside the skip-gram model, word embeddings can be learned using the continuous bag-of-words model (CBOW) but also by other models such as described, e.g., in Collobert and Weston (2008), Mnih and Hinton (2009), and Turian et al. (2010).

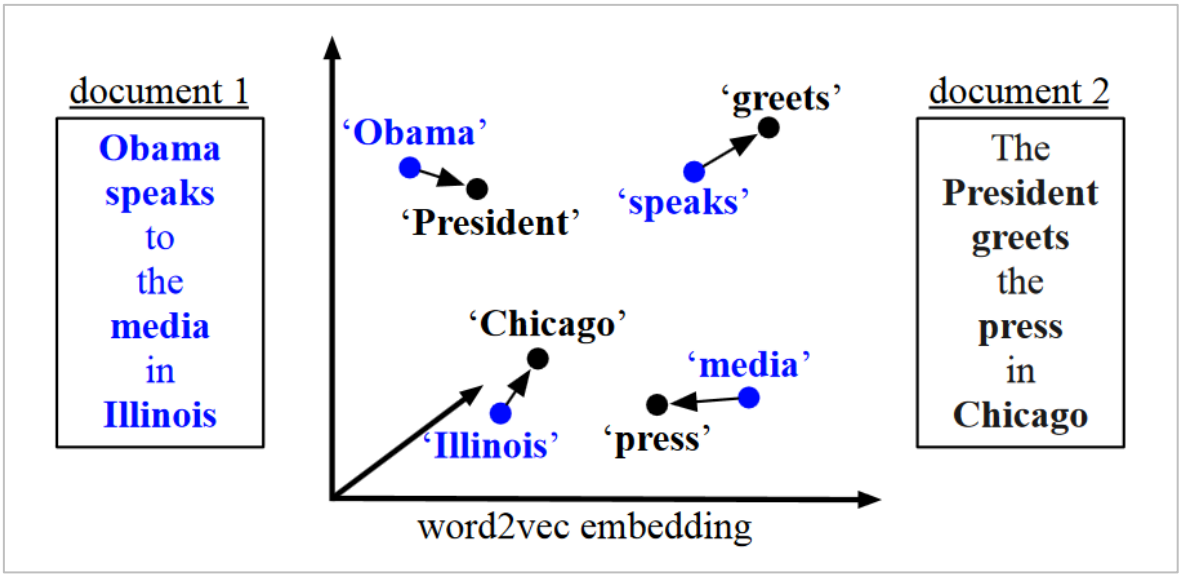

Figure 13: The idea underlying word mover's distance model. *Note*: From From Word Embeddings to Document Distances (Kusner et al. 2015: 957).

In the present study, learning the word embeddings of our news corpus was done by word2vec combined with the skip-gram model following Kusner et al. (2015) for which

purpose, we used the free gensim[88] Python library provided by Radim Řehůřek. Word2vec is an unsupervised method and was therefore directly applied to our news corpus without any previous training. In order to achieve a more effective learning of word embeddings, the model was additionally run on the Wiki news corpus and the Reuters corpus.

The Wikinews corpus used is a Wikipedia dump database file of 41.9 MB released under the Creative Commons Attribution Share-Alike and was downloaded on March 31, 2016, from the Wikimedia downloads site.[89] The data, which include news articles, templates, media/file descriptions, and primary meta-pages is originally stored in the file in XML format. The news articles included are from different categories, including sports, politics, health, etc., published on the Wikinews site[90] during the time period from November 13, 2004, to January 14, 2016. To extract and clean the news texts, including the headlines, from the XML file, a WikiExtractor 2.5 script was applied, written in Python by Giuseppe Attardi and Antonio Fuschetto. WikiExtractor is part of a free software Tanl released under the terms of the GNU General Public License 3. The corpus contains a total of 8,375,718 tokens and 124,890 types. In turn, the Reuters corpus contains 10,788 news texts which amount to 1.3 million words. The Reuters corpus was accessed within the NLTK Python library.

Before application of the word2vec, the data was preprocessed by following the procedure described in the previous subsection. Thus, the preprocessing included tokenization, sentence splitting, lemmatization, and POS tagging. The latter was applied in order to extract only words with high semantic content, such as nouns, and including open compounds, verbs, and adjectives. After learning the word embeddings in our news corpus, WMD was applied. The WMD script[91] used was developed by Matt Kusner.

Finally, in the submodule *Filtering*, the pairs of thematically related sentences obtained in the previous submodule were further analyzed in order to filter out those that do not fulfill the further conditions necessary for contradiction. Thus, after the sentences were aligned, a simplified temporal processing was done, including identifying the explicit temporal expressions and events but excluding determining the temporal relations between the sentences within a text. Also a simplified identification of co-referent expressions was conducted. For aligning the sentences, word2vec combined with a continuous skip-gram model was applied as previously mentioned. For the purpose of time processing, a SuTime library for recognizing and normalizing time expressions was applied, which is a component of the

[88] https://radimrehurek.com/gensim/
[89] https://dumps.wikimedia.org/
[90] https://de.wikinews.org/wiki/Hauptseite
[91] https://github.com/mkusner/wmd

Stanford CoreNLP. The SuTime library was accessed via a SuTime Python wrapper[92] provided by Frank Blechschmidt.

7.2.3 Module Finding Contradictions

In the last step, the search for an occurrence of a contradiction cue identified in Chapter 5 was implemented. In the case, a contradiction cue was found, two sentences or parts of a single sentence were classified as a potential contradiction. To remind, borderline and fake contradictions were treated under the concept of a potential contradiction as well.

Thus, for finding the negated events, SentimentAnnotator was used, which is a component of the Stanford CoreNLP. Two aligned events were judged as contradictory or contrary if one was identified as negated or negative and the other as positive or asserted. For finding negation constructions frequently occurring in news texts such as, e.g., *no information on* (Section 5.2.1), a pre-defined list of word constructions was used.

For recognition whether an event mentioned in the news refers to a real or possible situation, an automatic event factuality annotation has been applied.[93] That is, each event was assigned with one of the four factuality categories following (Saurí and Pustejovsky 2012). These categories are *factual* indicating that the event corresponds to a fact in the world, *counterfactual* indicating that the event does not refer to any real situation in the world, *non-factual* indicating it is uncertain whether the event took place or will take place in the future and *underspecified* in case no information on status of the event is provided. The factual status of an event was determined by application of the vua-factuality, an online factuality profiler as a part of the NLP pipeline of the Newsreader project,[94] and was then assigned manually. The vua-factuality factuality profiler which was trained on the FactBank corpus annotated with factualities determines the factual status of the event based on three notions including the *certainty* (certain/probable/possible) of an event, *polarity* of an event (pos/neg) and *time* of an event (future/non-future).

The certainty of an event refers, on one side, to how certain it is that an event happened. In this context, certainty is described in terms of epistemic modality categories such as *certainty*, *possibility* and *probability*. On the other side, certainty is concerned with evidentiality, that is, the way the information about a situation (witnessed, inferred, seen, directly experienced, heard) was acquired (van Valin and LaPolla 1997; Aikhenvald 2004). Different types of evidence can impact the factuality evaluation of an event. Thus, something which is reported as seen is perceived more likely to be a fact than anything that was only inferred

[92] https://github.com/FraBle/python-sutime/blob/master/setup.py
[93] In some papers, also the terms factivity and veracity are used equivalently.
[94] http://ixa2.si.ehu.es/nrdemo/demo.php

(Saurí/Pustejovsky 2012).

The explicit markers of certainty are the verbal auxiliaries *must* and *have to*, the implicative verbs *to presume*, the adverbs *necessarily*, *certainly*, and the adjectives *certain* and *impossible*. In turn, the explicit markers of possibility are the verbal auxiliaries *may*, *can*, and *could*, the adverbs *maybe* and *perhaps*, and the adjective *possible*. Finally, the probability is expressed by the verbal auxiliaries *will* and *should*, the adverb *probably*, and the adjectives *likely* and *probable*. In the case the event could not be assigned to one of the three certainty values, it was identified to be uncertain. In turn, the time of the event was classified as non-future (for present and past events), future, and underspecified, in case the time reference could not be identified. Finally, the polarity of the event refers to whether the event is affirmed or negated. In the case the polarity of the event could not be identified, the event was classified as affirmed.

Thus, an event was classified as factual (or happened) if it was a certain event with positive polarity and a non-future time reference. In turn, an event was classified as counterfactual (not happened) if it was certain with negative polarity and non-future time reference. Finally, an event was classified as non-factual if it was not certain whether the event already happened or will happen in the future. In case, the factuality of the event could not be identified, it was marked as underspecified.

It is to add, that beside modality-denoting expressions, factuality profiling by using vua-factuality module relies also on other lexical cues of factuality which Saurí and Pustejovsky (2012) refer to as event-selecting predicates (ESP). ESPs which can be nouns, verbs and adjectives embed an event and by this means express the degree of its factuality. It is to note, that ESPs can be further divided into source-introducing predicates (SIP) and non-source introducing predicates (NSIP). Thus, the ESPs which mark a high degree of factuality belong e.g. implicative verbs (*to manage*), factive predicates (e.g. *to know*), perception verbs (e.g. *to see*), aspectual verbs (e.g. *to finish*, *to begin*, *to continue*, *to terminate*), change of state verbs (e.g. *to increase*, *to change*, *to improve*). In turn, a counterfactuality is expressed by the implicative verbs such as e.g. *to avoid* and *to prevent* (inherent negatives, s. Section 3.2.1). Events which are non-factual can be marked by the verbs of belief and opinion (*to think*, *to speculate*, *to suspect*, *to consider*, *to guess*, *to predict*, *to suggest*). Finally, the degree of factuality of an event is underspecified if embedded by the verbs of volition (*to want*, *to wish*, *to hope*), commissive (*to offer*), commitment (*to commit*, *to propose*), and inclination predicates (*willing*, *ready*, *eager*, *reluctant*).

For identifying the polarity, the sentiment analysis of the SentimentAnnotator from the Stanford CoreNLP was applied. Finally, the time of an event was identified by conducting a part-

of-speech tagging. The events introduced by ESPs were identified based on a pre-defined list of reporting expressions.

For detecting the contradictions arising from opposition meaning relation, the relations between the words were computed using the WordNet lexical resource. The WordNet was accessed within NLTK Python library.

For finding the numerical divergencies, the information of the part-of-speech tagging from the module *Preprocessing* for identifying numerals was used. In making the decision whether a contradiction is observed or not, additionally, a pre-defined list of ranges of numbers was applied. As mentioned above, in the module *Preprocessing*, all the numerals occurring in the texts have been normalized to a common comparable basis.

The implementation of detecting factual contradictions included only the detection of contradictions arising from the incorrect spelling of names. Incorporation of the world knowledge to determine factual contradictions stayed beyond the scope. Thus, for detecting divergencies in name spelling, simple edit-based, also known as string-based, similarity measures such as Levenshtein distance, Damerau-Levenshtein distance, and Jaro distance were applied on proper nouns. In particular, Sun, Ma, and Wang (2015), for instance, emphasized the efficiency of the Levenshtein distance for finding typographical errors in name spelling.

The general idea behind the edit-based metrics is to compare two strings (in our case, words) based on their individual characters and compute their similarity by counting the number of edits required to transform one string into the other. The edits include the steps of insertion, deletion, replacement, and matching. The metrics differ by giving a different cost to different types of edits or to different characters. Thus, e.g. Levenshtein distance computes the distance between two strings, s_1 and s_2, by assigning each editing step with the value 1 and measuring their distance as a minimum cost of edits need to transform string s_1 into string s_2. The similarity between the strings measured with Levenshtein metric is a subtraction of 1 and a normalized distance such as illustrated in (7.1):

$$(7.1) \quad sim_{Levenshtein}(s_1, s_2) = 1 - \frac{LevenshteinDist(s_1, s_2)}{\max(|s_1|, |s_2|)}$$

The Damerau-Levenshtein distance is similar to the Levenshtein distance but, in contrast to the latter, considers the transposed characters in the strings. Thus, in the case of *Jack* and *Jakc*, the number of edits amounts to 1 and similarity to 0.75 if computed using the Damerau-Levenshtein, while the number of edits computed by Levenshtein distance amounts to 2, and the similarity is 0.5, accordingly. In turn, the Jaro distance searches for common characters, counting the number of matching characters m and number of transpositions t (7.2):

(7.2) $sim_{Jaro}(s_1, s_2) = \frac{1}{3}\left(\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m}\right)$

For applying the metrics, the jellyfish Python module[95] provided by James Turk was used. The identification of proper nouns was conducted by applying POS tagging.

7.3 Results and Evaluation

The system was evaluated with regard to three aspects, which are (1) the overall performance of the system with and without applying the anaphora resolution (coreference resolution), (2) the performance of the system regarding the detection of each type of contradictions, and (3) the performance of the system regarding finding parts of contradiction. For evaluating the system, a Gold standard including 299 cases (290 pairs of sentences and 9 single sentences), which are contradictions of one type, and 12 cases (12 sentence pairs exclusively), which are contradictions of two types, was used. The dataset of the Gold standard contains 351 unique sentences in total with the possibility of constructing of 61776 potential contradictions and contrarieties of one type and of 123552 potential contradictions and contrarieties of two types.

The evaluation of the above-mentioned three aspects of system's performance was conducted by means of precision and recall evaluation metrics. While the recall shows how many relevant cases the system was able to find contained in a dataset, precision indicates how precise the system was in the finding of relevant cases. The recall was computed as the fraction of the number of relevant cases found to the number of all relevant cases in the dataset. In turn, precision was computed as the fraction of the number of relevant cases found to the number of all cases found. The system implementation was aimed at obtaining a high precision.

Thus, with regard to the overall performance (evaluation aspect 1), excluding the pre-step of the anaphora resolution, the system identified 297 cases (only sentence pairs) as (potential) contradictions/contrarieties from the dataset of 351 unique sentences. Of these, only 124 pairs were (potential) contradictions contained in the Gold standard, while 173[96] pairs were incorrectly identified as (potential) contradictions and 199 pairs incorrectly identified as non-contradiction (Table 30), indicating the system's precision of 0.417 (42%) and recall of 0.383 (38%) (Table 31). In turn, the system identified 343 cases as (potential) contradictions when the task of anaphora resolution was conducted previously. Of the 343 cases

[95] https://github.com/jamesturk/jellyfish

[96] Justifying the survey results on manually finding contradictions, we treated all cases that are not included in the dataset as incorrect. However, we cannot exclude that these cases include contradictions that have not been identified by the survey participants.

found, 176 cases were (potential) contradictions contained in the dataset, while 167 pairs did not constitute a contradiction (Table 30), indicating the system's precision of 0.513 (51%) and recall of 0.544 (54%) (Table 31).

	Contradiction	**Yes**	**No**	**Total**
Without anaphora resolution	Yes	124	199	323
	No	173	123056	123229
	Total	297	123255	123552
With anaphora Resolution	Yes	176	147	323
	No	167	123062	123229
	Total	343	123209	123552

Table 30: Confusion matrix for system's performance (in columns) compared to the Gold standard (in rows).

	Precision (%)	**Recall (%)**
Without anaphora resolution	42	38
With anaphora resolution	51	54

Table 31: The overall performance of the system evaluated using precision and recall.

Thus, the computed recall of 54% shows that the system was able to find more than the half of relevant sentence pairs, or contradictions contained in the dataset when the anaphora resolution was previously applied, in comparison to recall of 38% when the step of the anaphora resolution was left out. Also the precision of 51% indicates that the system was more precise in the detection of contradictions than in the case when the anaphora resolution was not previously applied where a precision of 42% was achieved. Therefore, when comparing the results of system's performance with and without previously applying the anaphora resolution, a slight improvement of the results was observed over when the anaphora resolution (coreference resolution) was conducted previously.

Regarding the precision of 51% and recall 54% the overall performance of the system (including anaphora resolution) can be evaluated as low. This can be explained by that not all proposed components of the system's modules could be completely implemented. This primarily refers to the detection of implicit contradictions. Also simplified negation and modality detection, and factuality profiling are to be mentioned in this context.

In addition to evaluation of overall performance of the system the system's performance was evaluated in the detection of each contradiction type with anaphora resolution previously conducted (evaluation aspect 2). It is to remind, that no sophisticated implementation for detecting implicit contradictions was made. Implementation of detection of factual contradictions which require world knowledge was completely left beyond the scope. The results are summarized in Table 32.

Thus, the results indicate that the system showed a better recall in the detection of explicit contradictions (0.890, or 89%) in comparison to recall in detection of implicit ones (0.321, or 32%). The system's precision by recognition of both explicit and implicit contradictions is almost similar (0.549, or 55% and 0.460, or 46%, respectively). Further, the system performed best in the detection of numerical contradictions, realized both explicitly and implicitly (precision of 0.611 (61%) and recall of 0.963 (96%) for explicit and precision of 0.685 (69%) and recall of 1.0 (100%) for implicit). This can be explained by observing that the detection of contradictions arising from numerical divergencies are less challenging than from other types. In contrast, the system failed in recognizing of implicit contradictions arising from negation, opposition and lexical divergencies. It is to remind, that no implementation of detection of factual contradictions was conducted.

Type of Contradiction	**Explicit-negation**	**Explicit-opposition**	**Explicit-numerical**	**Explicit-Lexical**	**Implicit-negation**	**Implicit-Opposition**	**Implicit-Numerical**	**Implicit-Lexical**	**Implicit-Factual**	**Total in dataset**
Explicit-negation	21	-	-	-	-	-	-	-		23
Explicit-opposition	-	7	-	-	-	-	-	-	-	16
Explicit-numerical	-	-	77	-	-	-	-	-	-	80
Explicit-lexical	-	-	-	-	-	-	-	-	-	8
Implicit- negation	-	-	-	-	-	-	-	-	-	37
Implicit-opposition	-	-	-	-	-	-	-	-	-	9
Implicit-numerical	-	-	-	-	-	-	63	-	-	92
Implicit-lexical	-	-	-	8	-	-	-	-	-	37
Implicit-factual	-	-	-	-	-	-	-	-	-	21
All found	45	17	126	18	27	9	93	8	0	

Table 32: Confusion matrix for contradiction types found by the system and contained in the dataset (Rows: Gold standard, Columns: Contradictio system).

Besides the overall system performance and its performance in finding contradictions of different types, the task of finding related units potentially containing a contradiction cue (evaluation aspect 3) from the module *Finding parts of contradiction* was evaluated. As mentioned in the previous subsection, we proposed the application of the WMD model, which is currently the preferred method for measuring textual similarity (dissimilarity in terms of the model). Thus, we evaluated whether the WMD can be applied in order to find the parts of a (potential) contradiction.

For the evaluation, a modified dataset of contradictions was used. That is, only 311 contradiction and contrariety cases served as the Gold standard, without the repetitions contained in the corpus. The dataset included in total 351 unique sentences with possibility of constructing of 61776 potential contradictions and contrarieties.

The raw results of the application of the WMD model are summarized in Table 33. Thus, in applying the WMD model, the system identified 295 pairs of sentences to be related or equivalent (in case of contradictions represented by one part). From these, 281 sentence pairs were contained in the corpus of contradictions, while 14 sentence pairs were not present there. According to these numbers, the computed precision and recall scores are 0.953 (95%) and 0.904 (90%) respectively indicating a high precision of the system in finding parts of potential contradictions and its good ability to recognize relevant cases in a dataset if the WMD model is applied.

Related	**Yes**	**No**	**Total**
Yes	281	30	311
No	14	61451	61465
Total	295	61481	61776

Table 33: Confusion matrix for the system performance in finding parts of contradiction (Gold standard in rows, system performance in columns).

8 Conclusions

The main purpose of news is to inform the reader about the current political, economic, and cultural events in the world. By that, the main requirements for the process of news production is an objective, uninvolved news reporting and an accurate, i.e. correct and consistent (contradiction-free) use of facts. A violation of the latter leads to the misinformation of the reader and, if detected, to a negative impact on the credibility and trustworthiness of the newspaper.

The recognition of contradictions in a (news) text is a challenging task for a human as it presupposes concentrated reading and requires world knowledge and the ability to analytically process the information obtained. Also, the age and mental capability of the reader plays an important role. Further, the task of contradiction recognition becomes even more difficult when dealing with contradictory facts occurring in texts that are separated by space and time. For this reason, the main aim of the present study was to propose a system for the automatic detection of contradictions occurring in news texts written in English.

The conceptual development of a CD system (conceptual design) was begun with the theoretical elaboration of the term contradiction and related aspects, by this means contributing to a better understanding of the concept. It is to be noted that also the related concept of contrariety was treated in the study along with that of contradiction.

First, the conditions have been defined that two sentences must satisfy in order to be judged as a contradiction. These conditions, underlying Aristotle's ontological view on contradiction, are, a) reference to the same thing, b) expression of the same proposition about this thing, c) the same time reference, d) negation as sentence operator, and e) exclusive and exhaustive disjunction. Also, Aristotle's doxastic view on contradiction was considered for the purpose of the study, according to which a contradiction arises in case something is believed to be and not to be. The semantic view on contradiction, in turn, was left beyond the scope as it is based on a truth value of a sentence which at present cannot be processed by a machine. This, in turn, poses a potential for a future work.

Second, it was observed that negation is not necessarily a cue of contradiction but can also be a signal of a related concept, which is contrariety. Moreover, it was shown that negation as a sentence operator can be realized in multiple ways and by different means and is not limited to the negation operator not.

Further, we showed that with regard to language phenomena such as vagueness, lexical ambiguity, and context of utterances, the presence or absence of a contradiction in speech and text is a matter of interpretation. Here, the terms, borderline contradiction and fake contradiction were used to address these issues.

Finally, we addressed the status of contradictions with regard to modality and presupposition. In case of modality, there are still discussions on the question of when two propositions, of which one or both contain expressions of modality, are to be regarded as contradictory. In dealing with modal propositions and contradiction in the present study, we followed the approach of computational linguistics proposed in de Marneffe et al. (2008) and considered the sentences with opposite modalities/factuality to be contradictory, such as with possible and not_possible, actual and not_actual, and necessary and not_necessary. In contrast, the status of contradiction in the context of presupposition, either truth-based or pragmatic, was only theoretically discussed and not further considered for the system. The concept of potential contradiction has been introduced to justify the limitations of the CD system under vagueness, ambiguity, context, presupposition, and modality.

In turn, the practical elaboration of the concept of contradiction with regard to news texts included identifying the contradiction cues and describing the realization mechanisms leading to a contradiction. That is, in the empirical part of the study, it was found out that the realization of a disjunction as a necessary condition of contradiction is not limited to the negation operator *not*. Based on the compiled corpus of contradictions (and contrarieties) occurring in news texts, five types of contradiction cues were identified. Besides negation, these also include opposition relation, numerical, factual, and lexical divergencies. According to the type of relatedness of the parts, it was revealed that contradictions can be explicit (verbal similarity or equivalence of the contradiction parts) and implicit (no verbal similarity or equivalence of the contradiction parts; the carrier of implicit meaning are presupposition, entailment, and implicature).

Additionally, based on the evidence in the corpus, it was observed that contradictions of the kind, *Socrates is a man* and *Socrates is not a man*, which is an explicit contradiction according to the verbal correspondence of the parts with a negation as the sentence operator, rarely occur in news texts. In general, explicit contradictions realized by negation were observed to be rare in the compiled corpus of newswire and constitute only 7% of the whole corpus. Also, explicit-lexical (2%) and implicit-opposition (5%) contradictions were found to rarely occur in the corpus. In contrast, the explicit-numerical and implicit-numerical contradictions were observed to be the most frequent type, constituting 24.8% and 28.5% of the whole corpus, respectively. It is to be noted that the high frequency of numerical contradictions in the corpus can be explained by that they were easier to detect by the human survey participants than the contradictions arising from other cues. In general, the types of contradictions, according to their cues, are distributed as follows: negation (19%), opposition (8%), numerical (53%), lexical (14%), and factual (6%). In turn, explicit contradictions constitute 38.5% of the whole corpus, while implicit contradictions comprise 61.5%.

In contrast to the previous studies, our compiled corpus of contradictions occurring in news texts is the first corpus that provides pairs of contradictory sentences along with their co-text, contributing to a better data basis for the development and evaluation of CD systems. The corpus can also be used in the study of negation.

As the second step toward the conceptualization of the system (conceptual and physical design), the existing approaches and methods for contradiction detection were analyzed, and their weaknesses were identified. Thus, it was found, among others, that most existing systems deal only with the contradictions arising from negation and antonyms, which, in turn, does not cover all the possible varieties of contradictions. Moreover, the systems were tested on manually constructed data, without considering that the types of contradictions can vary for different text types. Further, the systems were tested on pairs of contradictory sentences taken isolated from their co-text, by this means losing valuable information for a more efficient contradiction detection. The task of the anaphora resolution is to be mentioned in this context. Finally, it was found that the systems do not consider interpreting contradictions under vagueness and contextual information.

Finally, in the third step, the main characteristics of the news texts have been summarized, by this means contributing to a better processing and interpretation of the textual data. It was shown that news texts are characterized by the use of standard language and rarely use figurative language. The frequent occurrence of reported speech, nominalizations, descriptive noun phrase, and a name noun phrase in the studies addressed were found to be typical for news texts as well. In contrast, modal verbs and contracted forms were observed by some studies to be less common for news texts. The latter explains the low frequency of negation contradictions as observed in the compiled corpus.

Following the elaboration of a conceptual design, a system architecture for an automatic CD task has been proposed, and the system has been implemented. In order for contradictions to be treated by the system under consideration of vagueness, ambiguity, modality, and context, the concept of potential contradiction has been introduced. The concept assumes that the final decision on contradiction is to be taken over by the user.

The results of the system's implementation showed that the models incorporating embedded text representations, using the WMD model as an example, perform well in the recognition of similar sentences in a corpus of texts, which potentially contain a contradiction cue. Moreover, the results showed that applying an anaphora resolution has a positive impact on finding contradictions and by this means improves the overall results of the CD system. Finally, the results showed that the implemented system performs best in the detection of numerical contradictions explicitly and implicitly occurring in news texts as well as of explicit contradictions arising from negation and opposition meaning relations. Also, the factual

contradictions arising from an incorrect name spelling can be accurately identified by the system.

The future work should be directed towards improving the system's functionality and extending its application space. Questions to be answered include: *Are there further types of contradictions in news texts that a system should be able to process? What is the best approach for dealing with modality?* Further questions are: *To what degree is the proposed system architecture applicable to other text types? Are there other types of contradictions typical for these text types? What are the cues and the realization mechanisms of these contradictions? Are additional system components required for detecting contradictions in other text types? If yes, what are the components that should be incorporated into the system? What is to be considered for processing the data in languages other than English? Are there other/additional components required?*

With concern to these questions, the future work must be done in three directions, which are further theoretical and practical elaborations on contradiction as well as extending the data basis and improving the general NLP tasks involved in the system.

That is, theoretical and practical elaboration on contradiction should include, among others, the study of contradiction types in newswire (including other multimedia used in the online news such as image and video) as well as the study of the particularities of contradiction realization in text types of other languages and the status of contradiction under modality.

An extension of the data basis should include a collection of contradictions occurring in newswire, texts of other types as well as other multimedia. By that, a particular focus on extending the data should be set on the development of a methodology for contradiction collection. The data basis should include contradictions along with their co-text and information on their context. Further, the work on the extension of the data should be directed to collecting the data in languages other than English.

Finally, future work should be done in improving the NLP tasks supporting contradiction detection, including anaphora resolution as well as negation, quantification, and modality processing, and textual entailment recognition, among others. Further, approaches should be developed for the detection of contradictions incorporating world knowledge. In this context, work should be also done in proposing approaches for revealing the information implicitly stated in the text. In this context, the identification and generation of presupposition, entailments, and implicatures, as well as logical inferencing, are a priority. Finally, the problem of identifying co-referent expressions (a reference to the same object in the real world) as well as determining the truth value of a sentence should be addressed, and appropriate approaches should be proposed.

Bibliography

Abbott, Barbara (2000): Presuppositions as Nonassertions. In *Journal of Pragmatics* 32 (10), pp. 1419–1437.

Abrusán, Márta (2011): Predicting the Presuppositions of Soft Triggers. In *Linguistics and Philosophy* 34 (6), pp. 491–535.

Abusch, Dorit (2002): Lexical Alternatives as a Source of Pragmatic Presuppositions. In *SALT* 12, p. 1.

Abusch, Dorit; Rooth, Mats (2004): Empty-domain Effects for Presuppositional and Non-presuppositional Determiners. In Hans Kamp, Barbara Hall Partee (Eds.): Context-dependence in the Analysis of Linguistic Meaning. 1st ed. Amsterdam, London: Elsevier, pp. 7–27.

Agerri, Rodrigo (2008): Metaphor in Textual Entailment. In: Proceedings of the 22nd International Conference on Computational Linguistics, Posters (COLING 2008). Manchester, UK, August 18-22, pp. 2–6.

Agichtein, Eugene; Askew, Walt; Liu, Yandong (2008): Combining Lexical, Syntactic, and Semantic Evidence for Textual Entailment Classification. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Agirre, Eneko; Banea, Carmen; Cardie, Claire; Cer, Daniel; Diab, Mona; Gonzalez-Agirre, Aitor et al. (2015): SemEval-2015 Task2: Semantic Textual Similarity, English, Spanish and Pilot on Interpretability. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). Denver, CO, USA, June 4-5, pp. 252–263.

Agirre, Eneko; Banea, Carmen; Cardie, Claire; Cer, Daniel; Diab, Mona; Gonzalez-Agirre, Aitor et al. (2014): SemEval-2014 Task 10: Multilingual Semantic Textual Similarity. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). Dublin, Ireland, August 23-24, pp. 81–91.

Agirre, Eneko; Banea, Carmen; Cer, Daniel; Diab, Mona; Gonzalez-Agirre, Aitor; Mihalcea, Rada et al. (2016): SemEval-2016 Task 1: Semantic Textual Similarity, Monolingual and Cross-Lingual Evaluation. In: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016). San Diego, CA, USA, June 16-17, pp. 497–511.

Agirre, Eneko; Cer, Daniel; Diab, Mona; Gonzalez-Agirre, Aitor (2012): SemEval-2012 Task 6: A Pilot on Semantic Textual Similarity. In: Proceedings of the First Joint Conference on Lexical and Computational Semantics (SEM 2012). Montreal, Canada, June 7-8, pp. 385–393.

Agirre, Eneko; Cer, Daniel; Diab, Mona; Gonzalez-Agirre, Aitor; Guo, Weiwei (2013): *SEM 2013 shared task: Semantic Textual Similarity. In: Proceedings of the Second Joint Conference on Lexical and Computational Semantics, Volume 1: Proceedings of the Main Conference and the Shared Task (SEM 2013), June 13-14. Atlanta, Georgia, pp. 32–43.

Agirre, Eneko; Lopez de Lacalle, Oier (2018): Clustering WordNet Word Senses. In Jeremy King, Sandro Sessarego (Eds.): Language Variation and Contact-Induced Change, vol. 260. Amsterdam: John Benjamins Publishing Company, pp. 121–130.

Aikhenvald, Alexandra (2004): Evidentiality. Oxford, New York: Oxford University Press.

Aitchison, Jean (2007): The Word Weavers. Newshounds and Wordsmiths. Cambridge: Cambridge University Press.

Albrecht, Jason E.; O'Brien, Edward J. (1993): Updating a Mental Model. Maintaining Both Local and Global Coherence. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19 (5), pp. 1061–1070.

Alexandersson, Jan; Reithinger, Norbert; Maier, Elisabeth (1997): Insights into the Dialogue Processing of VERBMOBIL. In: Proceedings of the fifth conference on Applied Natural Language Processing (ANLC 1997). Washington, DC, March 31-April 3. Morristown, NJ, USA: Association for Computational Linguistics, pp. 33–40.

Allan, Stuart (2006): Online News. Journalism and the Internet. Maidenhead: Open University Press.

Allen, James F. (1983): Maintaining Knowledge about Temporal Intervals. In *Communications of the ACM* 26 (11), pp. 832–843.

Allen, James F. (1991): Temporal Reasoning and Planning. In James F. Allen, Henry A. Kautz, Richard N. Pelavin, Tenenberg. Josh D. (Eds.): Reasoning about Plans. San Mateo, California, USA: Morgan Kaufmann Publishers, Inc., pp. 1–68.

Allen, James F. (1995): Natural language understanding. 2. ed., 3rd print. Redwood City, Calif.: Benjamin/Cummings.

Allwood, Jens S.; Andersson, Lars-Gunnar; Dahl, Östen (1977): Logic in Linguistics. Cambridge: Cambridge University Press.

Alxatib, Sam; Pagin, Peter; Sauerland, Uli (2013): Acceptable Contradictions. Pragmatics or Semantics? A Reply to Cobreros et al. In *Journal of Philosophical Logic* 42 (4), pp. 619–634.

Andrade, Daniel; Tsuchida, Masaaki; Onishi, Takashi; Ishikawa, Kai (2013): Detecting Contradiction in Text by Using Lexical Mismatch and Structural Similarity. In: Proceedings of the 10th National Institute of Informatics Testbeds and Community for Information access Research Conference (NTCIR 2013). Tokyo, Japan, June 18-21.

Aone, Chinatsu; Bennett, Scott William (1995): Evaluating Automated and Manual Acquisition of Anaphora Resolution Strategies. In: Proceedings of the 33rd Annual Meeting on Association for Computational Linguistics (ACL 1995). Cambridge, MA, USA, June 26-30. Morristown, NJ, USA: Association for Computational Linguistics, pp. 122–129.

Apel, Max; Ludz, Peter (1976): Philosophisches Wörterbuch. 6. Aufl. Berlin: Walter de Gruyter.

Aristoteles (2000): On Sophistical Refutations. Cambridge, Mass.: Harvard University Press [u.a.].

Aristotle (199-?): Prior Analytics. Raleigh, N.C, Boulder, Colo: Alex Catalogue; NetLibrary.

Aristotle; Ackrill, John L. (1994): Categories and De Interpretatione. Reprint. Oxford: Clarendon Press.

Aristotle; Barnes, Jonathan (1993): Posterior Analytics. 2 Rev.ed.

Aristotle; Kirwan, Christopher (1993): Metaphysics. Books [gamma], [delta], and [epsilon]. 2. ed. Oxford: Clarendon Press.

Asher, Nicholas (1993): Reference to Abstract Objects in Discourse. Dordrecht: Springer Netherlands.

Asher, Nicholas; Hardt, Daniel; Busquets, Joan (2001): Discourse Parallelism, Ellipsis, and Ambiguity. In *Journal of Semantics* 18 (1), pp. 1–25.

Asher, Nicholas; Lascarides, Alex (2003): Logics of Conversation. Cambridge: Cambridge University Press.

Atlas, Jay D.; Levinson, Stephen C. (1981): It-clefts, Informativeness and Logical Form. Radical pragmatics. In Peter Cole (Ed.): Radical Pragmatics. New York: Academic Press, pp. 32–37.

Attali, Yigal; Burstein, Jill (2006): Automated Essay Scoring with E-Rater V.2. In *Journal of Technology, Learning, and Assessment* 4 (3).

Audi, Robert (2015): The Cambridge Dictionary of Philosophy. Third edition. New York City: Cambridge University Press.

August, Diane L.; Flavell, John H.; Clift, Renee (1984): Comparison of Comprehension Monitoring of Skilled and Less Skilled Readers. In *Reading Research Quarterly* 20 (1), p. 39.

Austin, John Langshaw (1962): How to Do Things with Words. Cambridge Mass.: Harvard University Press.

Averbuch, Mordechai; Karson, Tom H.; Ben-Ami, Benjamin; Maimon, Oded; Rokach, Lior (2004): Context-Sensitive Medical Information Retrieval. In *Studies in health technology and informatics* 107 Pt 1, pp. 282–286.

Averintseva-Klisch, Maria (2013): Textkohärenz. Heidelberg: Winter.

Baeza-Yates, Ricardo; Ribeiro, Berthier de Araújo Neto (2011): Modern Information Retrieval. The Concepts and Technology behind Search. Second edition. Harlow, London, New York, Boston, San Francisco, Toronto, Sydney, Singapore, Hong Kong, Tokyo, Seoul, Taipei, New Delhi, Cape Town, Madrid, Mexico City, Amsterdam, Munich, Paris, Milan: Addison-Wesley/Pearson.

Baines, Paul (2007): Daniel Defoe. Robinson Crusoe - Moll Flanders. Basingstoke: Palgrave Macmillan.

Baker, Chris (1969): Double Negatives. In *Paper in Linguistics* 1 (1), pp. 16–40.

Baker, Collin; Ellsworth, Michael; Erk, Katrin (2007): SemEval'07 Task 19: Frame Semantic Structure Extraction. In: Proceedings of the 4th International Workshop on Semantic Evaluation (SemEval 2007). Prague, Czech Republic, June 23-24, pp. 99–104.

Baker, Kathy; Bloodgood, Michael; Dorr, Bonnie; Filardo, Nathaniel; Levin, Lori; Piatko, Christine (2010): A Modality Lexicon and its Use in Automatic Tagging. In: Proceedings of the Seventh Language Resources and Evaluation Conference (LREC 2010). Valletta, Malta, May 17-23.

Baker, Linda (1985): How Do We Know When We Don't Understand? Standards for Evaluating Text Comprehension. In Donna-Lynn. Forrest-Pressley, G. E. MacKinnon, Gary T. Waller (Eds.): Metacognition, Cognition, and Human Performance. New York: Academic Press.

Baker, Linda; Anderson, Richard I. (1982): Effects of Inconsistent Information on Text Processing. Evidence for Comprehension Monitoring. In *Reading Research Quarterly* 17 (2), p. 281.

Baker, Linda; Zimlin, Laurie (1989): Instructional Effects on Children's Use of Two Levels of Standards for Evaluating Their Comprehension. In *Journal of Educational Psychology* 81 (3), pp. 340–346.

Balduccini, Marcello; Baral, Chitta; Lierler, Yuliya (2007): Knowledge Representation and Question Answering. In Bruce Porter, Vladimir Lifschitz, Frank van Harmelen (Eds.): Handbook of Knowledge Representation, vol. 3. Amsterdam, Oxford: Elsevier, pp. 779–819.

Baldwin, Breck (1997): CogNIAC. In: Proceedings of a Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts (ANARESOLUTION 1997). Madrid, Spain, July 11. Morristown, NJ, USA: Association for Computational Linguistics, pp. 38–45.

Baldwin, Breck; Daynidhi, Krishna (2014): Natural Language Processing with Java and LingPipe Cookbook. Over 60 Effective Recipes to Develop Your Natural Language Processing (NLP) Skills Quickly and Effectively. Birmingham: Packt Publishing.

Banfield, Ann (2014): Unspeakable Sentences (Routledge Revivals). Narration and Representation in the Language of Fiction.

Banko, Michele; Cafarella, Michael; Soderland, Stephen; Broadhead, Matthew; Etzioni, Oren (2007): Open Information Extraction from the Web. In: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007). Hyderabad, India, January 6-12.

Banko, Michele; Etzioni, Oren (2008): The Tradeoffs Between Open and Traditional Relation Extraction. In: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL 2008). Columbus, OH, USA, June 15-20.

Bär, Daniel; Zesch, Torsten; Gurevych, Irina (2013): DKPro Similarity: An Open Source Framework for Text Similarity. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013). Sofia, Bulgaria, August 4-9, pp. 121–126.

Bar-Haim, Roy; Dagan, Ido; Dolan, Bill; Ferro, Lisa; Giampiccolo, Danilo (2006): The Second PASCAL Recognising Textual Entailment Challenge. In: Proceedings of the Second PASCAL Challenges Workshop on Recognising Textual Entailment. Venice, Italy, April 10.

Barnwell, Katharine Grace Lowry (1969): A Grammatical Description of Mbembe (Adun Dialect). A cross river language. London: Dept. of General Linguistics Univ. College.

Bartsch, Renate; Vennemann, Theo (1972): Semantic Structures. A Study in the Relation between Semantics and Syntax. Frankfurt/M.: Athenäum Verlag.

Bateman, John A. (2008): Multimodality and Genre. A Foundation for the Systematic Analysis of Multimodal Documents. Basingstoke, Hampshire: Palgrave Macmillan.

Beall, J. C.; Colyvan, Mark (2001): Looking for Contradictions. In *Australasian Journal of Philosophy* 79 (4), pp. 564–569.

Bean, David L.; Riloff, Ellen (1999): Corpus-based Identification of Non-anaphoric Noun Phrases. In: Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics on Computational Linguistics (ACL 1999). College Park, MD, USA, June 20-26. Morristown, NJ, USA: Association for Computational Linguistics, pp. 373–380.

Beaugrande, Robert-Alain de; Dressler, Wolfgang Ulrich (1981): Introduction to Text Linguistics. London, New York: Longman.

Beaver, David; Condoravdi, Cleo (2003): A Uniform Analysis of 'Before' and 'After'. In *SALT* 13, p. 37.

Beaver, David I. (2001): Presupposition and Assertion in Dynamic Semantics. Stanford Calif., Dordrecht?: CSLI Publications; FoLLI.

Beaver, David I.; Clark, Brady Z. (2008): Sense and Sensitivity. How Focus Determines Meaning. Oxford: Blackwell.

Beaver, David I.; Geurts, Bart (2014): Presupposition. In Edward N. Zalta (Ed.): The Stanford Encyclopedia of Philosophy. Winter 2014: Metaphysics Research Lab, Stanford University.

Becker, Jörg; Kuropka, Dominik (2003): Topic-based Vector Space Model. In: Proceedings of the 6th International Conference on Business Information Systems. Colorado, CO, USA, June 4-6, pp. 7–12.

Becker-Donner, Etta (1965): Die Sprache der Mano. Graz: Böhlau.

Bednarek, Monika (2006): Epistemological Positioning and Evidentiality in English News Discourse. A Text-driven Approach. In *Text & Talk - An Interdisciplinary Journal of Language, Discourse Communication Studies* 26 (6), pp. 635–660.

Bednarek, Monika (2008): Emotion Talk across Corpora. Basingstoke England, New York: Palgrave Macmillan.

Bednarek, Monika; Caple, Helen (2012): News Discourse. London: Continuum International Publishing.

Bell, Allan (1991): The Language of News Media. Oxford, UK: Blackwell.

Bennett, Jonathan (1988): Events and Their Names. Oxford: Clarendon.

Bentivogli, Luisa; Clark, Peter; Dagan, Ido; Giampiccolo, Danilo (2009): B. The Fifth Pascal Recognizing Textual Entailment Challenge. In: Proceedings of the Second Text Analysis Conference (TAC 2009). Gaithersburg, MD, USA, November 16-17: NIST.

Bentivogli, Luisa; Clark, Peter; Dagan, Ido; Giampiccolo, Danilo (2010): The Sixth PASCAL Recognizing Textual Entailment Challenge. In: Proceedings of the Third Text Analysis Conference (TAC 2010). Gaithersburg, MD, USA, November 15-16.

Bentivogli, Luisa; Clark, Peter; Dagan, Ido; Giampiccolo, Danilo (2011): The Seventh PASCAL Recognizing Textual Entailment Challenge. In: Proceedings of the Text Analysis Conference (TAC 2011). Gaithersburg, MD, USA, November 14-15.

Bergen, Anke von; Bergen, Karl von (1993): Negative Polarität im Englischen. Tübingen: Narr.

Bergler, Sabine (2005): Conveying Attitude with Reported Speech. In James G. Shanahan, Yan Qu, Janyce Wiebe (Eds.): Computing Attitude and Affect in Text: Theory and Applications, vol. 20. Berlin/Heidelberg: Springer Science & Business Media, pp. 11–22.

Besten, Hans den (1986): Double Negation and the Genesis of Afrikaans. In Pieter Muysken, Norval Smith (Eds.): Substrata versus Universals in Creole Genesis. Papers from the Amsterdam Creole Workshop, April 1985. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 185–230.

Béziau. Jean-Yves (2006): Paraconsistent Logic! In *Sorites* 17, pp. 17–26.

Bhagat, Rahul; Hovy, Eduard (2013): What Is a Paraphrase? In *Computational Linguistics* 39 (3), pp. 463–472.

Biber, Douglas; Conrad, Susan (2009): Register, Genre, and Style. Cambridge: Cambridge University Press.

Biber, Douglas; Johannson, Stig; Leech, Geoffrey; Conrad, Susan; Finegan, Edward (1999): Longman Grammar of Spoken and Written English. 2. impr. Harlow: Longman.

Biega, Joanna; Kuzey, Erdal; Suchanek, Fabian M. (2013): Inside YAGO2s: a Transparent Information Extraction Architecture. In: Proceedings of the 22nd International Conference on World Wide Web Companion (WWW 2013). Rio de Janeiro, Brazil. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee (ACM Digital Library), pp. 325–328.

Bird, Steven; Klein, Ewan; Loper, Edward (2009): Natural Language Processing with Python. Sebastopol: O'Reilly Media Inc.

Björkelund, Anders; Farkas, Richárd (2012): Data-driven Multilingual Coreference Resolution Using Resolver Stacking. In: Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012). Jeju Island, Korea, July 12-14, pp. 49–55.

Björkelund, Anders; Kuhn, Jonas (2012): Phrase Structures and Dependencies for End-to-End Coreference Resolution. In: Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012). Mumbai, India, December, pp. 145–154.

Björkelund, Anders; Kuhn, Jonas (2014): Learning Structured Perceptrons for Coreference Resolution with Latent Antecedents and Non-local Features. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguist (ACL 2014). Baltimore, MD, USA, June 23-25, pp. 47–57.

Blackburn, Simon (2016): The Oxford Dictionary of Philosophy. Third edition. Oxford: Oxford University Press.

Blei, David M.; Ng, Andrew Y.; Jordan, Michael (2003): Latent Dirichlet Allocation. In *The Journal of Machine Learning Research* 3, 993-1022.

Bloor, Thomas; Bloor, A. Meriel (1995): The Functional Analysis of English. A Hallidayan Approach. London: Edward Arnold.

Bobrow, Danny G.; Condoravdi, Cleo; Crouch, Dick; Paiva, Valeria de; Karttunen, Lauri; King, Tracy H. et al. (2007): Precision-focused Textual Inference. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29. Morristown, NJ, USA: Association for Computational Linguistics, pp. 16–21.

Bolinger, Dwight (1960): Linguistic Science and Linguistic Engineering. In *WORD* 16 (3), pp. 374–391.

Bond, Oliver (2016): Negation Through Reduplication and Tone. Implications for the Lexical Functional Grammar/Paradigm Function Morphology Interface. In *Journal of Linguistics* 52 (02), pp. 277–310.

Bonevac, Daniel (1987): Deduction. Introductory Symbolic Logic. London: Mayfield / Eurospan.

Bonini, Adair (2009): The Distinction Between News and Reportage in the Brazilian Journalistic Context: A Matter of Degree. In Charles Bazerman, Adair Bonini, Débora de Carvalho Figueiredo (Eds.): Genre in a Changing World. Fort Collins Colo., West Lafayette Ind: WAC Clearinghouse; Parlor Press, pp. 196–222.

Bonini, Nicolao; Osherson, Daniel; Viale, Riccardo; Williamson, Timothy (1999): On the Psychology of Vague Predicates. In *Mind and Language* 14 (4), pp. 377–393.

Borkin, Ann (1971): Polarity Items in Questions. In : Papers from the 5th Regional Meeting of the Chicago Linguistic Society. Chicago, IL, USA, April 18–19. University of Chicago: Chicago Linguistic Society, pp. 53–62.

Bos, Johan; Markert, Katja (2005): Recognising Textual Entailment with Logical Inference. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT-EMNLP 2005). Vancouver, Canada, October 6-8. Morristown, NJ, USA: Association for Computational Linguistics, pp. 628–635.

Bowman, Samuel R.; Angeli, Gabor; Potts, Christopher; Manning, Christopher D. (2015): A Large Annotated Corpus for Learning Natural Language Inference. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015). Lisbon, Portugal, September 1. Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 632–642.

Boyd, Adriane; Gegg-Harrison, Whitney; Byron, Donna (2005): Identifying Non-referential It. In: Proceedings of the ACL Workshop on Feature Engineering for Machine Learning in Natural Language Processing (FeatureEng 2005). Ann Arbor, MI, USA, June 29. Morristown, NJ, USA: Association for Computational Linguistics, p. 40.

Brants, Thorsten (2000): TnT: a Statistical Part-of-speech Tagger. In: Proceedings of the Sixth Conference on Applied Natural Language Processing (ANLC 2000). Seattle, Washington. Morristown, NJ, USA: Association for Computational Linguistics, pp. 224–231.

Breck, Eric (2009): A Simple System for Detecting Non-entailment. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Breiman, Leo (2001): Random Forests. In *Machine Learning* 45 (1), pp. 5–32.

Brennan, Susan E.; Friedman, Marilyn W.; Pollard, Carl J. (1987): A Centering Approach to Pronouns. In: Proceedings of the 25th Annual Meeting on Association for Computational Linguistics (ACL 1987). Stanford, California, July 6-9. Morristown, NJ, USA: Association for Computational Linguistics, pp. 155–162.

Brighton, Paul; Foy, Dennis (2007): News Values. London: SAGE.

Briscoe, Ted; Carroll, John (1997): Automatic Extraction of Subcategorization from Corpora. In: Proceedings of the fifth conference on Applied Natural Language Processing (ANLC 1997). Washington, DC, March 31-April 3. Morristown, NJ, USA: Association for Computational Linguistics, pp. 356–363.

Broderick, James F.; Miller, Darren W. (2007): Consider the Source. A Critical Guide to 100 Prominent News and Information Sites on the Web. Medford N.J.: CyberAge Books/Information Today.

Brody, Bruch (1967): Logical Terms, Glossary of. In Paul Edwards (Ed.): The Encyclopedia of Philosophy. New York: Macmillan.

Brown, Gillian; Yule, George (1991): Discourse Analysis. Cambridge: Cambridge University Press.

Büring, Daniel (1997): The Meaning of Topic and Focus. The 59th Street Bridge Accent. London: Routledge.

Büring, Daniel; Hartmann, Katharina (2001): The Syntax and Semantics of Focus-sensitive Particles in German. In *Natural Language and Linguistic Theory* 19 (2), pp. 229–281.

Busemann, Stephan; Declerck, Thierry; Diagne, Abdel Kader; Dini, Luca; Klein, Judith; Schmeier, Sven (1997): Natural Language Dialogue Service for Appointment Scheduling Agents. In: Proceedings of the fifth conference on Applied Natural Language Processing (ANLC 1997). Washington, DC, March 31-April 3. Morristown, NJ, USA: Association for Computational Linguistics, pp. 25–32.

Bussmann, Hadumod (2006): Routledge Dictionary of Language and Linguistics. London: Routledge.

Bybee, Joan (1985): Morphology. A Study of the Relation between Meaning and Form. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Byron, Donna K. (2001): Resolving Pronominal Reference to Abstract Entities. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL 2002). Philadelphia, PA, USA, July 7-12. Morristown, NJ, USA: Association for Computational Linguistics, p. 80.

Caldas-Coulthard, Carmen R. (1994): On Reporting Reporting: the Representation of Speech in Factual and Factional Narratives. In Malcolm Coulthard (Ed.): Advances in Written Text Analysis. London: Routledge, pp. 295–308.

Cambria, Erik; Speer, Robert; Havasi, Catherine; Hussain, Amir (2010): SenticNet. A Publicly Available Semantic Resource for Opinion Mining. In: Proceedings of the Conference on Artificial Intelligence (AAAI 2010).

Campion, Nicolas; Martins, Daniel; Wilhelm, Alice (2009): Contradictions and Predictions. Two Sources of Uncertainty That Raise the Cognitive Interest of Readers. In *Discourse Processes* 46 (4), pp. 341–368.

Capell, Arthur; Hinch, H. E. (1970): Maung Grammar. Texts and Vocabulary. The Hague: Mouton.

Caple, Helen (2010): Doubling-up: Allusion and Bonding in Multi-semiotic News Stories. In Monika Bednarek, J. R. Martin (Eds.): New Discourse on Language. Functional Perspectives on Multimodality, Identity, and Affiliation. London, New York: Continuum, pp. 111–133.

Cappon, René J. (1982): The word. An Associated Press Guide to Good News Writing. New York: Associated Press.

Cardie, Claire (1992): Learning to Disambiguate Relative Pronouns. In: Proceedings of the Tenth National Conference on Artificial intelligence (AAAI 1992). San Jose, CA, USA, July 12-16, pp. 38–43.

Carletta, Jean (1996): Assessing Agreement on Classification Tasks: the Kappa Statistic. In *Computational Linguistics* 22 (2), pp. 249–254.

Carreras, Xavier; Màrquez, Lluís (2004): Introduction to the CoNLL-2004 Shared Task: Semantic Role Labeling. In: Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL 2004). Boston, MA, USA, May 6-7, pp. 89–97.

Carreras, Xavier; Màrquez, Lluís (2005): Introduction to the CoNLL-2005 Shared Task: Semantic Role Labeling. In: Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL 2005). Ann Arbor, MI, USA, June 29-30, pp. 152–164.

Carreras, Xavier; Màrquez, Lluís; Padró, Lluís (2002): Named Entity Extraction Using AdaBoost. In: Proceedings of the 6th Conference on Natural language Learning (CoNLL 2002). Taipei, Taiwan, August 26-30. Morristown, NJ, USA: Association for Computational Linguistics, pp. 1–4.

Carstensen, Kai-Uwe; Ebert, Christian; Ebert, Cornelia; Jekat, Susanne; Klabunde, Ralf; Langer, Hagen (Eds.) (2010): Computerlinguistik und Sprachtechnologie. Eine Einführung. 3., überarb. und erw. Aufl. Heidelberg: Spektrum Akad. Verlag.

Castillo, Julio J. (2009): Sagan in TAC2009: Using Support Vector Machines in Recognizing Textual Entailment and TE Search Pilot task. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Castillo, Julio J.; Alonso i Alemany, Laura (2008): An approach using Named Entities for Recognizing Textual Entailment. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Chambers, Nathanael; Manning, Christopher D.; Cer, Daniel; Grenager, Trond; Hall, David; Kiddon, Chloe et al. (2007): Learning Alignments and Leveraging Natural Logic. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29. Morristown, NJ, USA: Association for Computational Linguistics, p. 165.

Chan, Lorna K.S.; Cole, Peter G.; Barfett, Shirley (1987): Comprehension Monitoring. Detection and Identification of Text Inconsistencies by LD and Normal Students. In *Learning Disability Quarterly* 10 (2), pp. 114–124.

Chang, Angel X.; Manning, Christopher D. (2012): SUTime: A Library for Recognizing and Normalizing Time Expressions. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012). Istanbul, Turkey, May.

Chapman, W. W.; Bridewell, W.; Hanbury, P.; Cooper, G. F.; Buchanan, B. G. (2001): A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries. In *Journal of biomedical informatics* 34 (5), pp. 301–310.

Charniak, Eugene (2000): A Maximum-Entropy-Inspired Parser. In: Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference (NAACL 2000). Seattle, WA, USA, April 29-May 04, pp. 132–139.

Chen, Minmin; Xu, Zhixiang; Weinberger, Kilian Q.; Sha, Fei (2012): Marginalized Denoising Autoencoders for Domain Adaptation. In: Proceedings of the 29th International Coference on International Conference on Machine Learning. (ICML 2012). Edinburgh, Scotland, June 26-July 01, pp. 1627–1634.

Chierchia, Gennaro; McConnell-Ginet, Sally (1990): Meaning and Grammar. An Introduction to Semantics. Cambridge Mass.: MIT Press.

Chklovski, Timothy; Pantel, Patrick (2004): VerbOcean. Mining the Web for Fine-Grained Semantic Verb Relations.

Chung, Sandra; Timberlake, Alan (1985): Tense, Aspect and Mood. In Timothy Shopen (Ed.): Language Typology and Syntactic Description: Volume 1. Clause Structure. Cambridge: Cambridge University Press, pp. 241–258.

Church, Kenneth Ward (1988): A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text. In: Proceedings of the Second Conference on Applied Natural Language Processing (ANLC 1988). Austin, TX, USA, February 9-12. Morristown, NJ, USA: Association for Computational Linguistics, p. 136.

Ciravegna, Fabio; Lavelli, Alberto; Gilardoni, Luca; Matiasek, Johannes; Mana, Nadia; Mazza, Silvia et al. (1999): FACILE. Classifying Texts Integrating Pattern Matching and Information Extraction Intelligence. In: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI 1999). Stockholm, Sweden, July 31-August 6: Morgan Kaufmann, pp. 890–897.

Clark, Herbert H. (1974): Semantics and Comprehension. In Thomas A. Sebeok (Ed.): Linguistics and Adjacent Arts and Sciences. The Hague: Mouton, pp. 1291–1428.

Clark, Herbert H. (1977): Bridging. In Philip Nicholas Johnson-Laird, P. C. Wason (Eds.): Thinking. Readings in Cognitive Science. Reprinted. Cambridge: Cambridge University Press, pp. 243–263.

Clark, Herbert H.; Clark, Eve V. (1977): Psychology and Language. An Introduction to Psycholinguistics. New York, NY: Harcourt Brace Jovanovich.

Clark, Kevin; Manning, Christopher D. (2015): Entity-Centric Coreference Resolution with Model Stacking. In: Proceedings of the Association for Computational Linguistics Conference (ACL 2015).

Clark, Kevin; Manning, Christopher D. (2016): Deep Reinforcement Learning for Mention-Ranking Coreference Models. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing.

Clark, Peter; Harrison, Philip (2008): Recognizing Textual Entailment with Logical Inference. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Clark, Peter; Harrison, Philip (2009): An Inference-Based Approach to Recognizing Entailment. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Clark, Peter; Murray, William R.; Thompson, John; Harrison, Philip; Hobbs, Jerry R.; Fellbaum, Christiane (2007): On the role of lexical and world knowledge in RTE3. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29, pp. 54–59.

Cobreros, Pablo; Egré, Paul; Ripley, David; van Rooij, Robert (2010): Tolerant, Classical, Strict. In *Journal of Philosophical Logic* 41 (2), pp. 347–385.

Cohen, Jacob (1960): A Coefficient of Agreement for Nominal Scales. In *Educational and Psychological Measurement* 20 (1), pp. 37–46.

Cohen, Shai (2009): On the Semantics of Too and Only: Distinctness and Subsumption. Ph.D. Thesis. Amherst, MA: University of Massachussets Amherst.

Cohen, William W.; Ravikumar, Pradeep; Fienberg, Stephen E. (2003): A Comparison of String Distance Metrics for Name-matching Tasks. In: Proceedings of IJCAI-03 Workshop on Information Integration on the Web (IIWeb 2003), vol. 2003. Acapulco, Mexico, August 9-10.

Collins, Michael (2003): Head-Driven Statistical Models for Natural Language Parsing. In *Computational Linguistics* 29 (4), pp. 589–637.

Collobert, Ronan; Weston, Jason (2008): A Unified Architecture for Natural Language Processing. In: Proceedings of the 25th International Conference on Machine Learning (ICML 2008). Helsinki, Finland, July 5-9. Helsinki, Finland: University of Helsinki], pp. 160–167.

Collobert, Ronan; Weston, Jason; Bottou, Léon; Karlen, Michael; Kavukcuoglu, Koray; Kuksa, Pavel (2011): Natural Language Processing (Almost) from Scratch. In *Journal of Machine Learning Research* 12, pp. 2493–2537.

Conboy, Martin (2007): The Language of the News. London: Routledge.

Condoravdi, Cleo; Crouch, Dick; Paiva, Valeria de; Stolle, Reinhard; Bobrow, Daniel G. (2003): Entailment, Intensionality and Text Understanding. In: Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies 2003 (NAACL-HLT 2003). Edmonton, Canada, May 27-June 1. Morristown, NJ, USA: Association for Computational Linguistics, pp. 38–45.

Conger, Anthony J. (1980): Integration and Generalization of Kappas for Multiple Raters. In *Psychological Bulletin* 88 (2), pp. 322–328.

Cook, Roy T. (2009): A Dictionary of Philosophical Logic. Edinburgh: Edinburgh University Press.

Cooper, Robin (1983): Quantification and Syntactic Theory. Dordrecht: Springer Netherlands.

Cooper, Robin; Crouch, Dick; van Eijck, Jan; Fox, Chris; van Genabith, Josef; Jaspars, Jan et al. (1996): Using the Framework.

Cotter, Colleen (2010): News Talk. Investigating the Language of Journalism. Cambridge: Cambridge University Press.

Councill, Isaac G.; McDonald, Ryan; Velikovich, Leonid (2010): What's Great and What's Not: Learning to Classify the Scope of Negation for Improved Sentiment Analysis. In: Proceedings of the Workshop on Negation and Speculation in Natural Language Processing (NeSp-NLP 2010). Uppsala, Sweden, July 10, pp. 51–59.

Cramer, Peter A. (2011): Controversy as News Discourse. Dordrecht: Springer.

Crammer, Koby; Singer, Yoram (2001): Ultraconservative Online Algorithms for Multiclass Problems. In G. Goos, J. Hartmanis, J. van Leeuwen, David Helmbold, Bob Williamson (Eds.): Computational Learning Theory, vol. 2111. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 99–115.

Crystal, David (2008): Language and the Internet. 2nd ed., reprinted. Cambridge: Cambridge University Press.

Cyffer, Norbert; Ebermann, Erwin; Ziegelmeyer, Georg (Eds.) (2009): Negation Patterns in West African Languages and Beyond. Amsterdam: John Benjamins Publishing Company.

Dagan, Ido (2013): Recognizing Textual Entailment. Models and Applications. San Rafael, Calif.: Morgan & Claypool Publ.

Dagan, Ido; Dolan, Bill; Magnini, Bernardo; Roth, Dan (2009): Recognizing Textual Entailment. Rational, Evaluation and Approaches. In *Nat. Lang. Eng.*

Dahl, Östen (2010b): Typology of Negation. In Laurence R. Horn (Ed.): The Expression of Negation. Berlin: De Gruyter Mouton, pp. 9–38.

Dahl, Östen (1979): Typology of Sentence Negation. In *Linguistics* 17 (1-2), pp. 79–106.

Dahlgren, Kathy (1974): The Pragmatics of Presuppostion. In George Bedell (Ed.): UCLA Papers in Syntax 5. Los Angeles, USA: University of California, Dept. of Linguistics, pp. 1–16.

D'Andrea, Alessia; Ferri, Fernando; Grifoni, Patrizia; Guzzo, Tiziana (2015): Approaches, Tools and Applications for Sentiment Analysis Implementation. In *International Journal of Computer Applications* 125 (3), pp. 26–33.

Davidson, Donald (1967): The Logical Form of Action Sentences. In Nicholas Rescher (Ed.): The Logic of Decision and Action. Pittsburgh, USA: University of Pittsburgh Press.

Davies, Mark; Fleiss, Joseph L. (1982): Measuring Agreement for Multinomial Data. In *Biometrics* 38 (4), p. 1047.

Deerwester, Scott; Dumais, Susan T.; Furnas, George W.; Landauer, Thomas K.; Harshman, Richard (1990): Indexing by Latent Semantic Analysis. In *Journal of the American Society for Information Science* 41 (6), pp. 391–407.

Defoe, Daniel (2001): Robinson Crusoe. London: Electric Book Co.

Delin, Judy (1992): Properties of It -Cleft Presupposition. In *Journal of Semantics* 9 (4), pp. 289–306.

Delin, Judy (1995): Presupposition and Shared Knowledge in It-Clefts. In *Language and Cognitive Processes* 10 (2), pp. 97–120.

Delmonte, Rodolfo; Tonelli, Sara; Tripodi, Rocco (2009): Semantic Processing for Text Entailment with VENSES. In: Proceedings of the Second Text Analysis Conference (TAC 2009). Gaithersburg, MD, USA, November 16-17: NIST.

Dempster, Arthur; M. Laird, Nan; Rubin, Donald B. (1977): Maximum Likelihood from Incomplete Data via the EM algorithm. In *Journal of the Royal Statistical Society. Series B (Methodological)* 39.

Denecke, Kerstin; Brosowski, Marko (2010): Topic Detection in Noisy Data Sources. In: Proceedings of the Fifth International Conference on Digital Information Management (ICDIM 2010). Thunder Bay, Canada, July 5-8: IEEE, pp. 50–55.

Dînşoreanu, Mihaela; Potolea, Rodica (2013): A Scalable Approach for Contradiction Detection Driven by Opinion Mining. In: Proceedings of International Conference on Information Integration and Web-based Applications & Services (IIWAS 2013). Vienna, Austria, December 2-4. New York, New York, USA: ACM Press, pp. 7–15.

Dobrin, Lisa M.; Nichols, Lynn; Rodriguez, Rosa M. (Eds.) (1991): Papers from the 27th Regional Meeting of the Chicago Linguistic Society 27 (2). Chicago, Ill.: Chicago Linguistic Society.

Dolan, Bill; Quirk, Chris; Brockett, Chris (2004): Unsupervised Construction of Large Paraphrase Corpora. In: Proceedings of the 20th international conference on Computational Linguistics (COLING 2004). Geneva, Switzerland, August 23-27. Morristown, NJ, USA: Association for Computational Linguistics, Article Nr. 350.

Downing, Angela; Locke, Philip (2006): English Grammar. A University Course. London: Routledge.

Dowty, David (1991): Thematic Proto-roles and Argument Selection. In *Language* 67 (3), pp. 547–619.

Dryer, Matthew S. (1988): Universals of Negative Position. In Michael Hammond, Edith A. Moravcsik, Jessica Wirth (Eds.): Studies in Syntactic Typology, vol. 17. Amsterdam: John Benjamins Publishing Company, p. 93.

Dryer, Matthew S. (2005): Negative Morphemes. In Martin Haspelmath, Matthew S. Dryer, David Gil, Bernard Comrie (Eds.): World Atlas of Language Structures. Oxford: Oxford University Press, pp. 454–457.

Dryer, Matthew S. (2009): Verb-Object-Negative Order in Central Africa. In Norbert Cyffer, Erwin Ebermann, Georg Ziegelmeyer (Eds.): Negation Patterns in West African Languages and Beyond, vol. 87. Amsterdam: John Benjamins Publishing Company, pp. 307–362.

Elbourne, Paul D. (2008): Demonstratives as Individual Concepts. In *Linguistics and Philosophy* 31 (4), pp. 409–466.

Elbourne, Paul D. (2005): Situations and Individuals. Cambridge, Mass., London: MIT Press.

Ennals, Rob; Trushkowsky, Beth; Agosta, John Mark (2010): Highlighting Disputed Claims on the Web. In: Proceedings of the 19th International Conference on World Wide Web (WWW 2010). Raleigh, NC, USA, April 26-30. New York, New York, USA: ACM Press, pp. 341–350.

Erk, Katrin; Padó, Sebastian (2006): Shalmaneser - A Toolchain For Shallow Semantic Parsing. In: Proceedings of International Conference on Language Resources and Evaluation (LREC 2006). Genoa, Italy, May 22-28.

Esuli, Andrea; Sebastiani, Fabrizio (2006): SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. In: Proceedings of International Conference on Language Resources and Evaluation (LREC 2006). Genoa, Italy, May 22-28, pp. 417–422.

Evans, Richard (2000): A Comparison of Rule-Based and Machine Learning Methods for Identifying Non-nominal It. In G. Goos, J. Hartmanis, J. van Leeuwen, Dimitris N. Christodoulakis (Eds.): NLP 2000 - Proceedings of the International Conference on Natural Language Processing, vol. 1835. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 233–240.

Fader, Anthony; Soderland, Stephen; Etzioni, Oren (2011): Identifying Relations for Open Information Extraction. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2011). Edinburgh, United Kingdom, July 27-31, pp. 1535–1545.

Fairclough, Norman (1995): Media Discourse. London: Arnold.

Farkas, Donka F.; Bruce, Kim B. (2010): On Reacting to Assertions and Polar Questions. In *Journal of Semantics* 27 (1), pp. 81–118.

Feez, Susan; Iedema, Rick; White, Peter; Silva Joyce, Helen de (2008): Media Literacy. Surry Hills, N.S.W.: NSW Adult Migrant Education Service.

Fellbaum, Christiane (Ed.) (1998): WordNet. An Electronic Lexical Database. Cambridge, Mass: MIT Press.

Fenton, Natalie (2012): News in the Digital Age. In Stuart Allan (Ed.): The Routledge Companion to News and Journalism. Rev. ed. London u.a.: Routledge, pp. 557–567.

Ferrández, Óscar; Munoz, Rafael; Palomar, Manuel (2009): Alicante University at TAC 2009: Experiments in RTE. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Ferro, Lisa; Gerber, Laurie; Mani, Inderjeet; Sundheim, Beth (2005): TIDES 2005: Standard for the Annotation of Temporal Expressions. Technical Report. The MITRE Corporation.

Filatova, Elena; Hatzivassiloglou, Vasileios (2003): Domain-Independent Detection, Extraction, and Labeling of Atomic Events. In: Proceedings of the International Conference in Recent Advances in Natural Language Processing (RANLP 2003). Borovetz, Bulgaria, September 10-12.

Fillmore, Charles J. (1968): The Case for Case. In Emmon W. Bach, Robert T. Harms, Charles J. Fillmore (Eds.): Universals in Linguistic Theory. London: Holt Rinehart & Winston, pp. 1–25.

Fillmore, Charles J. (1975): Types of Lexical Information. In F. Kiefer (Ed.): Studies in Syntax and Semantics. Dordrecht: Springer Netherlands, pp. 109–137.

Fillmore, Charles J. (1976): Frame Semantics and the Nature of Language. In *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280 (1), pp. 20–32.

Fillmore, Charles J.; Ruppenhofer, Josef; Baker, Collin F. (2004): FrameNet and Representing the Link between Semantic and Syntactic Relations. In Chu-ren Huang, Winifred Lenders (Eds.): Computational Linguistics and Beyond. Taipei: Institute of Linguistics, Academia Sinica, pp. 19–62.

Fine, Kit (1975): Vagueness, Truth and Logic. In *Synthese* 30 (3-4), pp. 265–300.

Finkel, Jenny Rose; Grenager, Trond; Manning, Christopher (2005): Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. In: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (ACL 2005). Ann Arbor, Michigan. Morristown, NJ, USA: Association for Computational Linguistics, pp. 363–370.

Fintel, Kai von (1999): NPI Licensing, Strawson Entailment, and Context Dependency. In *Journal of Semantics* 16 (2), pp. 97–148.
Fintel, Kai von (2006): Modality and Language. In Donald M. Borchert (Ed.): Encyclopedia of Philosophy. 2nd ed. Detroit, London: Macmillan Reference USA.
Firth, John Rupert (1957): Papers in Linguistics. 1934 - 1951. London: Oxford University Press.
Fishman, Mark (1980): Manufacturing the News. Austin, Tex.: University of Texas Press.
Flavell, John H.; Speer, James Ramsey; Green, Frances L.; August, Diane L.; Whitehurst, Grover J. (1981): The Development of Comprehension Monitoring and Knowledge about Communication. In *Monographs of the Society for Research in Child Development* 46 (5), p. 1.
Fleiss, Joseph L. (1971): Measuring Nominal Scale Agreement among Many Raters. In *Psychological Bulletin* 76 (5), pp. 378–382.
Forbes, Graeme (1994): Modern Logic. A Text in Elementary Symbolic Logic. New York, NY: Oxford University Press.
Fowler, Roger (1991): Language in the News. Discourse and Ideology in the Press. 1st publ. London: Routledge.
Frege, Gottlob (2011): Über Sinn und Bedeutung (1892). In Günther Patzig (Ed.): Funktion, Begriff, Bedeutung. Fünf logische Studien. 1. Aufl., Textausg. Göttingen: Vandenhoeck & Ruprecht, pp. 23–46.
Freud, Sigmund (2009): Der Witz und seine Beziehung zum Unbewußten. Der Humor. 9., unveränd. Aufl. Frankfurt am Main: Fischer-Taschenbuch-Verlag.
Fries, Udo (1987): Summaries in Newspapers: A Textlinguistic Investigation. In Udo Fries (Ed.): The Structure of Texts. Tübingen: Narr, pp. 47–63.
Fundel, Katrin; Küffner, Robert; Zimmer, Ralf (2007): RelEx - Relation Extraction Using Dependency Parse Trees. In *Bioinformatics* 23 (3), pp. 365–371.
Gabrilovich, Evgeniy; Markovitch, Shaul (2006): Overcoming the Brittleness Bottleneck Using Wikipedia: Enhancing Text Categorization with Encyclopedic Knowledge. In: Proceedings of the Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference (AAAI 2006). Boston, MA, USA, July 16-20, pp. 1301–1306.
Galanis, Dimitrios; Malakasiotis, Prodromos (2008): AUEB at TAC 2008. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.
Galtung, Johan; Ruge, Mari Holmboe (2016): The Structure of Foreign News. In *Journal of Peace Research* 2 (1), pp. 64–90.
Ganeev, Bulat T. (2004): Protivorečija v jazyke i reči. Ufa: Baškirskij Gosudarstvennyj Pedagogičeskij Universitet.
Gans, Herbert J. (1979): Deciding What's News. A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time. 1. ed. New York: Pantheon Books.
Ganter, Viola; Strube, Michael (2009): Finding Hedges by Chasing Weasels. In: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers (ACL-IJCNLP 2009). Suntec, Singapore, August 4. Morristown, NJ, USA: Association for Computational Linguistics, p. 173.
Garner, Ruth (1980): Monitoring of Understanding. An Investigation of Good and Poor Readers' Awareness of Induced Miscomprehension of Text. In *Journal of Reading Behavior* 12 (1), pp. 55–63.
Garner, Ruth (1981): Monitoring of Passage Inconsistency among Poor Comprehenders. A Preliminary Test of the “Piecemeal Processing” Explanation. In *The Journal of Educational Research* 74 (3), pp. 159–162.
Garretson, Gregory; Ädel, Annelie (2008): Who's Speaking? Evidentiality in US Newspapers during the 2004 Presidential Campaign. In Annelie Ädel, Randi Reppen (Eds.): Corpora and Discourse. The Challenges of Different Settings. Amsterdam: Benjamins, pp. 157–188.
Garrison, Daniel H. (2004): The Student's Catullus. 3rd ed. Norman: University Of Oklahoma Press.
Gauker, Christopher (1998): What is a Context of Utterance? In *Philosophical Studies* 91 (2), pp. 149–172.
Gawron, Jean Mark (1996): Quantification, Quantificational Domains and Dynamic Logic. In Shalom Lappin (Ed.): The Handbook of Contemporary Semantic Theory. Oxford: Blackwell, pp. 247–267.
Geis, Michael L. (1987): The Language of Politics. New York, Berlin, Heidelberg, London, Paris, Tokyo: Springer.
Geurts, Bart; van der Sandt, Rob (2004): Interpreting Focus. In *Theoretical Linguistics* 30 (1), pp. 1–44.

Giampiccolo, Danilo; Dang, Hoa Trang; Magnini, Bernardo; Dagan, Ido (2008): The Fourth PASCAL Recognizing Textual Entailment Challenge. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Giampiccolo, Danilo; Magnini, Bernardo; Dagan, Ido; Dolan, Bill (2007): The Third PASCAL Recognizing Textual Entailment Challenge. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29. Morristown, NJ, USA: Association for Computational Linguistics, p. 1.

Giannakidou, Anastasia; Zeijlstra, Hedde (2017): The Landscape of Negative Dependencies: Negative Concord and N-words. In Martin Everaert, Henk van Riemsdijk (Eds.): The Companion to Syntax. Hoboken, NJ, USA: John Wiley & Sons, Inc, pp. 1–38.

Gildea, Daniel; Jurafsky, Daniel (2002): Automatic Labeling of Semantic Roles. In *Computational Linguistics* 28 (3), pp. 245–288.

Givón, Talmy (1979): On Understanding Grammar. New York: Academic Press.

Givón, Talmy (1984): Syntax. A Functional-typological Introduction. Amsterdam: Benjamins.

Givón, Talmy (1993): English Grammar. A Function-based Introduction. Amsterdam: Benjamins.

Glenberg, Arthur M.; Wilkinson, Alex Cherry; Epstein, William (1982): The Illusion of Knowing. Failure in the Self-assessment of Comprehension. In *Memory & Cognition* 10 (6), pp. 597–602.

Glinos, Demetrios G. (2008): Recognizing Textual Entailment at RTE4 with CERES. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Gödert, Winfried; Lepsky, Klaus; Nagelschmidt, Matthias (2012): Informationserschließung und Automatisches Indexieren. Ein Lehr- und Arbeitsbuch. Berlin: Springer.

Gomaa, Wael H.; Fahmy, Aly A. (2013): A Survey of Text Similarity Approaches. In *International Journal of Computer Applications* 68 (13), pp. 13–18.

Gonçalves, Pollyanna; Benevenuto, Fabrício; Cha, Meeyoung (2013): PANAS-t. A Pychometric Scale for Measuring Sentiments on Twitter.

Gottlieb, Paula (2015): Aristotle on Non-contradiction. In Edward N. Zalta (Ed.): The Stanford Encyclopedia of Philosophy. Summer 2015: Metaphysics Research Lab, Stanford University.

Govier, Trudy (2010): A Practical Study of Argument. 7th ed. Belmont, CA: Cengage Learning.

Graesser, Arthur C.; Kreuz, Roger J. (1993): A Theory of Inference Generation during Text Comprehension. In *Discourse Processes* 16 (1-2), pp. 145–160.

Graesser, Arthur C.; Singer, Murray; Trabasso, Tom (1994): Constructing Inferences during Narrative Text Comprehension. In *Psychological Review* 101 (3), pp. 371–395.

Green, Margaret M.; Igwe, G. E. (1963): A Descriptive Grammar of Igbo. Berlin: Akademie-Verlag.

Greenberg, Joseph Harold (1963): Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Joseph Harold Greenberg (Ed.): Universals of Language. Report of a Conference Held at Dobbs Ferry, N.Y., April 13-15, 1961. Cambridge (Mass.): MIT Press, pp. 73–113.

Grice, Herbert P. (1988): The Causal Theory of Perception. In Jonathan Dancy (Ed.): Aristotelian Society Supplementary Volume: Oxford University Press, pp. 121–168.

Grice, Herbert P. (1989): Studies in the Way of Words. Cambridge, Massachusetts., London, England: Harvard University Press.

Grim, Patrick (2004): What is a Contradiction? In Graham Priest, J. C. Beall, Bradley P. Armour-Garb (Eds.): The Law of Non-contradiction. Oxford: Clarendon, pp. 49–72.

Grosz, Barbara J.; Weinstein, Scott; Joshi, Aravind K. (1995): Centering: a Framework for Modeling the Local Coherence of Discourse. In *Computational Linguistics* 21 (2), pp. 203–225.

Gruber, Thomas R. (1993): A Translation Approach to Portable Ontology Specifications. In *Knowledge Acquisition* 5 (2), pp. 199–220.

Guarino, Nicola (1998): Formal Ontology and Information System. In: Proceedings of the Conference on Formal Ontology in Information Systems (FOIS 1998). Trento, Italy, June 6-8. Amsterdam, Netherlands: IOS Press, pp. 3–15.

Gunlogson, Christine (2001): True to Form: Rising and Falling Declaratives as Questions in English. Ph.D. Thesis. New York: Routledge.

Gunter, Barrie (2003): News and the Net. Mahwah, NJ: Lawrence Erlbaum.

Ha, Louisa; James, E. Lincoln (1998): Interactivity Reexamined. A Baseline Analysis of Early Business Web Sites. In *Journal of Broadcasting & Electronic Media* 42 (4), pp. 457–474.

Haack, Susan (1978): Philosophy of Logics. Cambridge: Cambridge University Press.

Hacioglu, Kadri; Chen, Ying; Douglas, Benjamin (2005): Automatic Time Expression Labeling for English and Chinese Text. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell et al. (Eds.): Computational Linguistics and Intelligent Text Processing, vol. 3406. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 548–559.

Hacioglu, Kadri; Pradhan, Sameer; Ward, Wayne H.; Martin, James H.; Jurafsky, Daniel (2004): Semantic Role Labeling by Tagging Syntactic Chunks. In: Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL 2004). Boston, MA, USA, May 6-7.

Haghighi, Aria; Klein, Dan (2009): Simple Coreference Resolution with Rich Syntactic and Semantic Features. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing Volume 3 (EMNLP 2009). Singapore, August 6-7. Morristown, NJ, USA: Association for Computational Linguistics, p. 1152.

Haghighi, Aria D.; Ng, Andrew Y.; Manning, Christopher D. (2005): Robust Textual Inference via Graph Matching. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT-EMNLP 2005). Vancouver, Canada, October 6-8. Morristown, NJ, USA: Association for Computational Linguistics, pp. 387–394.

Halliday, M. A. K.; Hasan, Ruqaiya (1976): Cohesion in English. London: Longman.

Hallin, Daniel C. (1987): Where? Cartography, Community and the Cold War. In Robert Karl Manoff, Michael Schudson (Eds.): Reading the News. New York, NY: Pantheon Books, pp. 109–145.

Han Ren, Donghong Ji; Wan, Jing (2009): WHU at TAC 2009: A Tri-categorization Approach to Textual Entailment Recognition. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Harabagiu, Sanda M.; Hickl, Andrew; Lacatusu, Finley V. (2006): Negation, Contrast and Contradiction in Text Processing. In: Proceedings of the Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference (AAAI 2006). Boston, MA, USA, July 16-20, pp. 755–762.

Hardt, Daniel (1997): An Empirical Approach to VP Ellipsis. In *Computational Linguistics* 23 (4), pp. 525–541.

Harris, Paul L.; Kruithof, Arne; Terwogt, Mark Meerum; Visser, Ton (1981): Children's Detection and Awareness of Textual Anomaly. In *Journal of Experimental Child Psychology* 31 (2), pp. 212–230.

Harris, Richard J.; Monaco, Gregory E. (1978): Psychology of Pragmatic Implication. Information Processing between the Lines. In *Journal of Experimental Psychology: General* 107 (1), pp. 1–22.

Harris, Zellig S. (1954): Distributional Structure. In *WORD* 10 (2-3), pp. 146–162.

Harris, Zellig S. (1981): Co-Occurrence and Transformation in Linguistic Structure. In Zellig S. Harris, Henry Hiż (Eds.): Papers on Syntax. Dordrecht: Springer Netherlands, pp. 143–210.

Harrison, Philip; Maxwell, Michael (1986): A New Implementation of GPSG. In: Proceedings of the 6th Canadian Conference on Artificial Intelligence (CSCSI 1986). Montreal, Canada, May 21-23, pp. 78–83.

Hartley, John (1982): Understanding News. London: Methuen.

Hashimoto, Chikara; Torisawa, Kentaro; Saeger, Stijn de; Oh, Jong-Hoon; Kazama, Jun'ichi (2012): Excitatory or Inhibitory: A New Semantic Orientation Extracts Contradiction and Causality from the Web. In: Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012). Jeju Island, Korea, July 12-14, pp. 619–630.

Haspelmath, Martin (1997): Indefinite Pronouns. Oxford: Clarendon Press.

Haspelmath, Martin (2005): Negative Indefinite Pronouns and Predicate Negation. In Martin Haspelmath, Matthew S. Dryer, David Gil, Bernard Comrie (Eds.): World Atlas of Language Structures. Oxford: Oxford University Press, pp. 466–469.

Hays, David. G. (1964): Dependency Theory. A Formalism and some Observations. In *Memorandum RM-. Rand Corporation* 42, pp. 1–39.

Heflin, Jeff; Pan, Zhengxiang (2004): A Model Theoretic Semantics for Ontology Versioning. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell et al. (Eds.): The Semantic Web – ISWC 2004, vol. 3298. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 62–76.

Heim, Irene (1982): On the Semantics of Definite and Indefinite Noun Phrases. Ph.D. Thesis: University of Massachusetts at Amherst.

Heim, Irene (1992): Presupposition Projection and the Semantics of Attitude Verbs. In *Journal of Semantics* 9 (3), pp. 183–221.

Heinämäki, Orvokki (1974): Semantics of English Temporal Connectives. Ph.D. Thesis: University of Indiana at Bloomington.

Heinemann, Margot; Heinemann, Wolfgang (2002): Grundlagen der Textlinguistik. Interaktion - Text - Diskurs. Berlin/Boston: De Gruyter.

Hickl, Andrew; Bensley, Jeremy (2007): A Discourse Commitment-based Framework for Recognizing Textual Entailment. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29. Morristown, NJ, USA: Association for Computational Linguistics, pp. 171–176.

Hickl, Andrew; Bensley, Jeremy; Williams, John; Roberts, Kirk; Rink, Bryan; Shi, Ying (2006): Recognizing Textual Entailment with Lcc's Groundhog System. In: Proceedings of the Second PASCAL Challenges Workshop on Recognising Textual Entailment. Venice, Italy, April 10, pp. 80–85.

Hidalgo Downing, Laura (2000): Negation, Text Worlds, and Discourse. The Pragmatics of Fiction. Stamford, Conn.: Ablex Pub. Corp.

Hirst, Graeme (1986): Semantic Interpretation and the Resolution of Ambiguity. Cambridge: Cambridge University Press.

Hirst, Graeme (1991): Existence Assumptions in Knowledge Representation. In *Artificial Intelligence* 49 (1-3), pp. 199–242.

Hobbs, Jerry R. (1976): Pronoun Resolution. New York, NY.

Hobbs, Jerry R. (1983): Why is Discourse Coherent. In Fritz Neubauer (Ed.): Coherence in Natural-language Texts. Hamburg: Buske, pp. 29–70.

Hobbs, Jerry R. (1985): On Coherence and Structure of Discourse. CSLI Technical Report. Stanford, CA, USA.

Hobbs, Jerry R.; Kehler, Andrew (1997): A Theory of Parallelism and the Case of VP Ellipsis. In: Proceedings of the 35th Annual Meeting on Association for Computational Linguistics. Madrid, Spain, July 7-12. Morristown, NJ, USA: Association for Computational Linguistics, pp. 394–401.

Hoffart, Johannes; Suchanek, Fabian M.; Berberich, Klaus; Lewis-Kelham, Edwin; Melo, Gerard de; Weikum, Gerhard (2011): YAGO2: Exploring and Querying World Knowledge in Time, Space, Context, and Many Languages. In: Proceedings of the 20st International World Wide Web Conference (WWW 2011). Hyderabad, India. New York, NY: ACM, pp. 229–233.

Hoffart, Johannes; Suchanek, Fabian M.; Berberich, Klaus; Weikum, Gerhard (2013): YAGO2. A Spatially and Temporally Enhanced Knowledge Base from Wikipedia. In *Artificial Intelligence* 194, pp. 28–61.

Hoffmann, Martin (2016): Die Onlinenachricht. In Dietz Schwiesau, Josef Ohler (Eds.): Nachrichten - klassisch und multimedial. Wiesbaden: Springer Fachmedien Wiesbaden, pp. 185–199.

Honeck, Richard P. (1971): A Study of Paraphrases. In *Journal of Verbal Learning and Verbal Behavior* 10 (4), pp. 367–381.

Hopper, Paul J.; Thompson, Sandra A. (1984): The Discourse Basis for Lexical Categories in Universal Grammar. In *Language* 60 (4), p. 703.

Horn, Laurence R. (2010b): Multiple Negation in English and Other Languages. In Laurence R. Horn (Ed.): The Expression of Negation. Berlin: De Gruyter Mouton, pp. 111–148.

Horn, Laurence R. (1978): Remarks on Neg-Raising. In Peter Cole (Ed.): Pragmatics. New York: Academic Press, pp. 129–217.

Horn, Laurence R. (1989): A Natural History of Negation. Chicago: University of Chicago Press.

Horn, Laurence R. (1991): Duplex Negatio Affirmat…: The Economy of Double Negation. In Lisa M. Dobrin, Lynn Nichols, Rosa M. Rodriguez (Eds.): Papers from the 27th Regional Meeting of the Chicago Linguistic Society 27 (2). Chicago, Ill.: Chicago Linguistic Society, pp. 78–106.

Horn, Laurence R. (2004): Implicature. In Laurence R. Horn, Gregory L. Ward (Eds.): The Handbook of Pragmatics. Malden MA: Blackwell Pub, pp. 3–28.

Horn, Laurence R. (2014): Contradiction. In Edward N. Zalta (Ed.): The Stanford Encyclopedia of Philosophy. Spring 2014: Metaphysics Research Lab, Stanford University.

Hoste, Veronique (2016): The Mention-Pair Model. In Massimo Poesio, Roland Stuckardt, Yannick Versley (Eds.): Anaphora Resolution. Algorithms, Resources, and Applications. New York NY: Springer Berlin Heidelberg, pp. 269–282.

Huang, Yang; Lowe, Henry J. (2007): A Novel Hybrid Approach to Automated Negation Detection in Clinical Radiology Reports. In *Journal of the American Medical Informatics Association* 14 (3), pp. 304–311.

Huddleston, Rodney D. (1984): Introduction to the Grammar of English. Cambridge: Cambridge University Press.

Hurford, James R.; Heasley, Brendan; Smith, Michael B. (2008): Semantics. A Coursebook. 2. ed., reprinted. Cambridge: Cambridge University Press.

Huttenlocher, Janellen; Higgins, E. Tory: Adjectives, Comparatives and Syllogisms. In *Psychological Review* 78, pp. 487–504.

Hyland, Ken (1998): Hedging in Scientific Research Articles. Amsterdam: John Benjamins Publishing Company.

Ifantidou, Elly (2009): Newspaper Headlines and Relevance. Ad Hoc Concepts in Ad Hoc Contexts. In *Journal of Pragmatics* 41 (4), pp. 699–720.

Iftene, Adrian (2008): UAIC Participation at RTE4. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Iftene, Adrian; Balahur-Dobrescu, Alexandra (2007): Hypothesis Transformation and Semantic Variability Rules Used in Recognizing Textual Entailment. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29. Morristown, NJ, USA: Association for Computational Linguistics, pp. 125–130.

Iftene, Adrian; Moruz, Mihai-Alex (2009): UAIC Participation at RTE5. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Ihlström Eriksson, Carina; Lundberg, Jonas (2004): A Genre Perspective on Online Newspaper Front Page Design. In *Journal of Web Engineering* 3.

Ivin, Aleksandr (2015): Iskusstvo Mislit Pravilno. Moskva: Prospekt.

Jackendoff, Ray S. (1987): Semantic Interpretation in Generative Grammar. Cambridge, Mass.: MIT Press.

Jackson, Peter; Moulinier, Isabelle (2002): Natural Language Processing for Online Applications. Text Retrieval, Extraction and Categorization. Amsterdam, Philadelphia, PA: John Benjamins Pub.

Jankowski, Nicholas W.; van Selm, Martine (2000): Traditional News Media on Line: An Examination of Added Values. In *Communications : The European Journal of Communication Research* 25 (1), pp. 85–101.

Jespersen, Otto (1917): Negation in English and Other Languages. København: Høst.

Jespersen, Otto (1924): The Philosophy of Grammar. S.l.: Allen and Unwin.

Jivani, Anjali Ganesh (2011): A Comparative Study of Stemming Algorithms. In *International Journal of Computer Applications in Technology* 2 (6), pp. 1930–1938.

Johnson-Laird, P. N.; Legrenzi, Paolo; Girotto, Vittorio (2004): How We Detect Logical Inconsistencies. In *Current Directions in Psychological Science* 13 (2), pp. 41–45.

Johnston, Jane L.; Forde, Susan: The Silent Partner: News Agencies and 21st Century News. In *International Journal of Communication* 5, pp. 195–214.

Jucker, Andreas H. (1992): Social Stylistics. Syntactic Variation in British Newspapers. Berlin, New York: Mouton de Gruyter.

Jurafsky, Dan; Martin, James H. (2000): Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Upper Saddle River, N.J.: Prentice Hall.

Jurafsky, Daniel; Martin, James H. (2008): Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 2. ed. Upper Saddle River, NJ: Prentice Hall.

Kahane, Howard (1995): Logic and Contemporary Rhetoric. The Use of Reason in Everyday Life. 7. ed. Belmont, Calif.: Wadsworth.

Kahrel, Pieter Johannus (1996): Aspects of Negation. Amsterdam, Netherlands: University of Amsterdam.

Kaiser, Elsi (2006): Negation and the Left Periphery in Finnish. In *Lingua* 116 (3), pp. 314–350.

Kalish, Donald; Montague, Richard; Mar, Gary (Eds.) (1980): Logic. Techniques of Formal Reasoning. Second edition. San Diego, London, Sydney: Harcourt Brace Jovanovich Publ.

Kamp, Hans (1981): A Theory of Truth and Semantic Representation. In Jeroen A. G. Groenendijk, Theo M.V. Janssen, Martin B.J. Stokhof (Eds.): Formal Methods in the Study of Language: University of Amsterdam, pp. 277–322.

Karttunen, Lauri (1971): Implicative Verbs. In *Language* 47 (2), p. 340.

Karttunen, Lauri (1973): Presuppositions of Compound Sentences. In *Linguistic Inquiry* 4 (2), pp. 169–193.

Karttunen, Lauri (1974): Presupposition and Linguistic Context. In *Theoretical Linguistics* 1 (1-3).

Karttunen, Lauri (1976): Discourse Referents. In James D. McCawley (Ed.): Syntax and Semantics. Notes from the Linguistic Underground. New York: Academic Press, pp. 363–385.

Karttunen, Lauri; Peters, Stanley (1979): Conventional Implicature. In Choon-Kyu Oh, David A. Dineen (Eds.): Presupposition. New York: Academic Press, pp. 1–56.

Katz, Jerrold J. (1972): Semantic Theory. New York, London: Harper and Row.

Kawahara, Daisuke; Inui, Kentaro; Kurohashi, Sadao (2010): Identifying Contradictory and Contrastive Relations between Statements to Outline Web Information on a Given Topic. In: Proceedings of the 23th International Conference on Computational Linguistics (COLING 2010). Beijing, China, August 23-27, pp. 534–542.

Keenan, Edward L. (1971): Two Kinds of Presupposition in Natural Language. In Charles J. Fillmore, D. Terence Langendoen (Eds.): Studies in Linguistic Semantics. New York, NY: Holt Rinehart and Winston, pp. 45–52.

Kehler, Andrew (2002): Coherence, Reference, and the Theory of Grammar. Stanford, Calif.: CSLI Publ.

Kempe, André (1993): A Probabilistic Tagger and an Analysis of Tagging Errors. Technical Report. Institut für maschinelle Sprachverarbeitung. Universität Stuttagart, Germany.

Kempson, Ruth M. (1975): Presupposition and the Delimitation of Semantics. Zugl.: London, Diss., 1973. Cambridge: Cambridge University Press.

Kennedy, Christopher; Boguraev, Branimir (1996): Anaphora for Everyone. In: Proceedings of the 16th Conference on Computational Linguistics (COLING 1996). Copenhagen, Denmark, August 5-9. Morristown, NJ, USA: Association for Computational Linguistics, p. 113.

Kenter, Tom; Rijke, Maarten de (2015): Short Text Similarity with Word Embeddings. In: Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM 2015). Melbourne, Australia, October 19-23. New York, New York: Assocation for Computing Machinery, pp. 1411–1420.

Khanmagomedov, B. G.-K. (1967): Tabasaraskij jazyk. In Evgenij Alekseevič Bokarev, Viktor V. Vinogradov (Eds.): Jazyki narodov SSSR - Iberijsko-kavkazskie jazyki. Moskva: Nauka, pp. 545–561.

Kibble, Rodger (2001): A Reformulation of Rule 2 of Centering Theory. In *Computational Linguistics* 27 (4), pp. 579–587.

Kilicoglu, Halil; Bergler, Sabine (2008): Recognizing Speculative Language in Biomedical Research Articles. A Linguistically Motivated Perspective. In *BMC bioinformatics* 9 Suppl 11, S10.

Kim, Hyun Duk; Zhai, ChengXiang (2009): Generating Comparative Summaries of Contradictory Opinions in Text. In: Proceedings of the ACM Eighteenth Conference on Information and Knowledge Management (CIKM 2009). Hong Kong, China, Novemver 2-6. [New York, N.Y.]: Association for Computing Machinery, pp. 385–394.

Kingsbury, Paul; Palmer, Martha; Marcus, Mitchell (2002): Adding Semantic Annotation to the Penn TreeBank. In: Proceedings of the Second International Conference on Human Language Technology Research (HLT 2002). San Diego, CA, USA, March 24-27, 252--256.

Kintsch, Walter (1988): The Role of Knowledge in Discourse Comprehension. A construction-integration model. In *Psychological Review* 95 (2), pp. 163–182.

Kintsch, Walter (1993): Information Accretion and Reduction in Text Processing. Inferences. In *Discourse Processes* 16 (1-2), pp. 193–202.

Kiparsky, Paul; Kiparsky, Carol (1970): Fact. In Manfred Bierwisch, Karl Erich Heidolph (Eds.): Progress in Linguistics. A Collection of Papers. Berlin/Boston: De Gruyter; De Gruyter Mouton, pp. 143–173.

Kipper, Karin; Dang, Hoa Trang; Palmer, Martha (2000): Class-Based Construction of a Verb Lexicon. In: Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Innovative Applications of Artificial Intelligence Conference (AAAI-2000 - IAAI-2000). Austin, TX, USA, July 30-August 3. Menlo Park, Calif.: AAAI Press; Cambridge.

Klavans, Judith L.; Chodorow, Martin (1992): Degrees of Stativity. In: Proceedings of the 14th Conference on Computational Linguistics (COLING 1992). Nantes, France. Morristown, NJ, USA: Association for Computational Linguistics, pp. 1126–1131.

Klein, Dan; Manning, Christopher D. (2003): Accurate Unlexicalized Parsing. In: Proceedings of the 41st Annual Meeting on Association for Computational Linguistics (ACL 2003). Sapporo, Japan, July 7-12. Morristown, NJ, USA: Association for Computational Linguistics, pp. 423–430.

Klima, Edward S. (1964): Negation in English. In Jerry A. Fodor, Jerrold J. Katz (Eds.): The Structure of Language. Readings in the Philosophy of Language. 6. print. Englewood Cliffs, N.J.: Prentice Hall, pp. 246–323.

Kloetzer, Julien; Saeger, Stijn de; Torisawa, Kentaro; Hashimoto, Chikara; Oh, Jong-Hoon; Sano, Motoki; Ohtake, Kiyonori (2013): Two-Stage Method for Large-Scale Acquisition of Contradiction Pattern Pairs using Entailment. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2013). Seattle, USA, October 18-21.

Kniffka, Hannes (1980): Soziolinguistik und empirische Textanalyse. Schlagzeilen- und Leadformulierung in amerikanischen Tageszeitungen. Berlin: Niemeyer.

Knox, John S. (2010): Online Newspapers: Evolving Genres and Evolving Theory. In Caroline Coffin, Theresa Lillis, Kieran O'Halloran (Eds.): Applied Linguistics Methods. A Reader. London: Routledge, pp. 33–51.

Kolhatkar, Varada; Hirst, Graeme (2012): Resolving "This-Issue" Anaphora. In: Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012). Jeju Island, Korea, July 12-14, pp. 1255–1265.

Kondakov, Nikolaj Ivanovič (1983): Wörterbuch der Logik. Hrsg. d. dt. Ausg. Erhard Albrecht. Leipzig: Bibliogr. Inst.

Koomen, Peter; Punyakanok, Vasin; Roth, Dan; Yih, Wen-tau (2005): Generalized Inference with Multiple Semantic Role Labeling Systems. In: Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL 2005). Ann Arbor, MI, USA, June 29-30, pp. 181–184.

Kotsonis, Miriam E.; Patterson, Charlotte J. (1980): Comprehension-Monitoring Skills in Learning-Disabled Children. In *Developmental Psychology* 16 (5), pp. 541–542.

Kouylekov, Milen; Negri, Matteo (2010): An Open-Source Package for Recognizing Textual Entailment. In: Proceedings of the ACL 2010 System Demonstrations (ACLDemos 2010). Uppsala, Sweden, July 13, pp. 42–47.

Krestel, Ralf; Bergler, Sabine; Witte, René (2008): A Belief Revision Approach to Textual Entailment Recognition. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Krestel, Ralf; Witte, René; Bergler, Sabine (2009): Believe It or Not: Solving the TAC 2009 Textual Entailment Tasks through an Artificial Believer System. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Krippendorff, Klaus (2012): Content Analysis. An Introduction to Its Methodology. 3rd ed. Thousand Oaks: SAGE Publications, Inc.

Krovetz, Robert (1993): Viewing Morphology as an Inference Process. In: Proceedings of the 16th Annual International Conference on Research and Development in Information Retrieval : Papers (SIGIR 1993). Pittsburgh, PA, USA. New York, New York, USA: ACM Press, pp. 191–202.

Krüger-Thielmann, Karin; Paijmans, Hans (2004): Informationserschließung. In Henning Lobin, Lothar Lemnitzer (Eds.): Texttechnologie. Perspektiven und Anwendungen. Tübingen: Stauffenburg, pp. 353–378.

Kudo, Taku; Matsumoto, Yuji (2004): A Boosting Algorithm for Classification of Semi-Structured Text. In: Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP 2004). Barcelona, Spain, July 25-26, pp. 301–308.

Kuipers, Aert H. (2015): The Squamish Language. Grammar, Texts, Dictionary. Berlin, Boston: De Gruyter Mouton.

Kusner, Matt J.; Sun, Yu; Kolkin, Nicholas I.; Weinberger, Kilian Q. (2015): From Word Embeddings to Document Distances. In: Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37 (ICML 2015). Lille, France, July 07-09, pp. 957–966.

La Roche, Walther von (2006): Einführung in den praktischen Journalismus. Mit genauer Beschreibung aller Ausbildungswege ; Deutschland, Österreich, Schweiz. 17., aktualisierte Aufl. Berlin: Econ.

Labov, William (1972): Negative Attraction and Negative Concord in English Grammar. In *Language* 48 (4), pp. 773–818.

Ladusaw, William A. (1980): Polarity Sensitivity as Inherent Scope Relations. Bloomington, Ind.: Indiana University Linguistics Club.

Lafferty, John D.; McCallum, Andrew; Pereira, Fernando C.N. (2001): Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In: Proceedings of the Eighteenth International Confernce on Machine Learning (ICML 2001). San Francisco, CA, USA, June 28-July 1, pp. 282–289.

Lage, Nilson (1979): Ideologia e técnica da notícia. Petrópolis: Editora Vozes.

Laka, Itziar (1990): Negation in Syntax. On the Nature of Functional Categories and Projections. Ph.D. dissertation. Cambridge, Mass: MIT Press.

Lakoff, George (1971): Presupposition and Relative Well-formedness. In Danny D. Steinberg, Leon A. Jakobovits (Eds.): Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology. Cambridge: Cambridge University Press, pp. 329–340.

Lakoff, George (1973): Hedges. A Study in Meaning Criteria and the Logic of Fuzzy Concepts. In *Journal of Philosophical Logic* 2 (4).

Lakoff, Robin (1971): Ifs, Ands, and Buts about Conjunction. In Charles J. Fillmore, D. Terence Langendoen (Eds.): Studies in Linguistic Semantics. New York, NY: Holt Rinehart and Winston, pp. 114–149.

Lamble, Stephen (2011): News as It Happens. An Introduction to Journalism. 1. publ. South Melbourne, Victoria: Oxford.

Landauer, Thomas; Dumais, Susan (2008): Latent Semantic Analysis. In *Scholarpedia* 3 (11), p. 4356.

Lapata, Mirella; Lascarides, Alex (2006): Learning Sentence-internal Temporal Relations. In *Journal of Artificial Intelligence Research* 27, pp. 85–117.

Lappin, Shalom; Leass, Herbert J. (1994): An Algorithm for Pronominal Anaphora Resolution. In *Computational Linguistics* 20 (4).

Lascarides, Alex; Asher, Nicholas (1993): Temporal Interpretation, Discourse Relations and Commonsense Entailment. In *Linguistics and Philosophy* 16 (5), pp. 437–493.

Lee, Heeyoung; Chang, Angel; Peirsman, Yves; Chambers, Nathanael; Surdeanu, Mihai; Jurafsky, Dan (2013): Deterministic Coreference Resolution Based on Entity-Centric, Precision-Ranked Rules. In *Computational Linguistics* 39 (4), pp. 885–916.

Lee, Heeyoung; Peirsman, Yves; Chang, Angel; Chambers, Nathanael; Surdeanu, Mihai; Jurafsky, Dan (2011): Stanford's Multi-Pass Sieve Coreference Resolution System at the CoNLL-2011 Shared Task. In: Proceedings of the Conference on Natural Language Learning Shared Task (CoNLL 2011).

Lehmann, Winfred Philipp (1974): Proto-Indo-European Syntax. Austin: University of Texas Press.

Leitner, Gerhard (1986): Reporting the "Events of the Day": Uses and Function of Reported Speech. In *Studia Anglica Posnaniensia* 18, pp. 189–204.

Lenat, Douglas B.; Guha, R. V. (1990): Building Large Knowledge-Based Systems. Representation and Inference in the CYC Project. Reading, Massachusetts: Addison-Wesley.

Lendvai, Piroska; Reichel, Uwe D. (2016): Contradiction Detection for Rumorous Claims. In: Proceedings of the Workshop Extra-Propositional Aspects of Meaning in Computational Linguistics (ExProM-2016). Osaka, Japan, December 12. Erscheinungsort nicht ermittelbar, Stroudsburg, PA, pp. 31–40.

Lennon, Paul (2004): Allusions in the Press. An Applied Linguistic Study. Berlin, New York: Mouton de Gruyter.

Levin, Beth (1993): English Verb Classes and Alternations. A Preliminary Investigation. Chicago, London: University of Chicago Press.

Levinson, Stephen C. (1983): Pragmatics. Cambridge: Cambridge University Press.

Lewis, David (1979): Possible Worlds. In Michael J. Loux (Ed.): The Possible and the Actual. Readings in the Metaphysica of Modality. Ithaca: Cornell University Press, pp. 182–189.

Li, Fangtao; Zheng, Zhicheng; Tang, Yang; Zhang, Xian H. (2008): THU QUANTA at TAC 2008 QA and RTE Track. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Light, Marc; Qiu, Xin Ying; Srinivasan, Padmini (2004): The Language of Bioscience. Facts, Speculations, and Statements in between. In: Proceedings of the HLT-NAACL 2004 Workshop: BioLINK 2004, Linking Biological Literature, Ontologies and Databases (HLT-NAACL 2004). Boston, Massachusetts, USA: Association for Computational Linguistics, pp. 17–24.

Light, Richard J. (1971): Measures of Response Agreement for Qualitative Data. Some Generalizations and Alternatives. In *Psychological Bulletin* 76 (5), pp. 365–377.

Lin, Chin-Yew (2004): ROUGE. A Package for Automatic Evaluation of Summaries. In : Text Summarization Branches Out: Proceedings of the ACL-04 Workshop (ACL 2004). Barcelona, Spain: Association for Computational Linguistics, pp. 74–81.

Lin, Dekang (1994): PRINCIPAR. In: Proceedings of the 15th conference on Computational linguistics (COLING 1994). Kyoto, Japan, August 5-9. Morristown, NJ, USA: Association for Computational Linguistics, p. 482.

Lin, Dekang; Pantel, Patrick (2001): DIRT - Discovery of Inference Rules from Text. In: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2001). San Francisco, CA, USA, August 26-29. New York, New York, USA: ACM Press, pp. 323–328.

Linebarger, Marcia C. (1987): Negative Polarity and Grammatical Representation. In *Linguistics and Philosophy* 10 (3), pp. 325–387.

Linebarger, Marcia Christine (1981): The Grammar of Negative Polarity. Bloomington: Indiana University Linguistics Club.

Linke, Angelika; Nussbaumer, Markus; Portmann, Paul R. (1994): Studienbuch Linguistik. 2. Auflage. Tübingen: Max Niemeyer Verlag.

Litkowski, Ken (2004): Senseval-3 Task. Automatic Labeling of Semantic Roles. In: Proceedings of SENSEVAL-3, the Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (SENSEVAL-3). Barcelona, Spain, July 25-26.

Liu, Bing (2015): Sentiment Analysis. Mining Opinions, Sentiments, and Emotions. Cambridge: Cambridge University Press.

Lobin, Henning (2010): Computerlinguistik und Texttechnologie. 1. Aufl. Paderborn: Fink.

Löbner, Sebastian (2013): Understanding Semantics. 2. ed. London: Routledge.

Long, Debra L.; Chong, Jennifer L. (2001): Comprehension Skill and Global Coherence. A Paradoxical Picture of Poor Comprehenders' Abilities. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27 (6), pp. 1424–1429.

Lorenz, Sven (1992): On the Role of Reasoning about Change in the Projection of Presuppositions. In *IWBS Report* 234.

Lovins, Julie B. (1968): Development of a Stemming Algorithm. In *Mechanical Translation and Computational Linguistics* 11 (1-2), pp. 22–31.

Lowth, Robert (1762): A Short Introduction to English Grammar. With Critical Notes. London: Hughs.

Lüger, Heinz-Helmut (1995): Pressesprache. Berlin: De Gruyter.

Łukasiewicz, Jan (1970): Selected Works. Edited by Ludwik Borkowski. Amsterdam: North-Holland Pub. Co.

Lund, Kevin; Burgess, Curt (1996): Producing High-dimensional Semantic Spaces from Lexical Co-occurrence. In *Behavior Research Methods, Instruments, & Computers* 28 (2), pp. 203–208.

Lyons, John (1977): Semantics. Volume 2. Cambridge: Cambridge University Press.

MacCartney, Bill (2009): Natural Language Inference. Ph.D. Thesis: Stanford University.

MacCartney, Bill; Galley, Michel; Manning, Christopher D. (2008): A Phrase-based Alignment Model for Natural Language Inference. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2008). Honolulu, HI, USA, October 25-27, pp. 802–811.

MacCartney, Bill; Grenager, Trond; Marneffe, Marie-Catherine de; Cer, Daniel; Manning, Christopher D. (2006): Learning to Recognize Features of Valid Textual Entailments. In: Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (HLT-EMNLP 2006). New York, NY, USA, June 4-9. Morristown, NJ, USA: Association for Computational Linguistics, pp. 41–48.

MacCartney, Bill; Manning, Christopher D. (2007): Natural Logic for Textual Inference. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29. Morristown, NJ, USA: Association for Computational Linguistics, pp. 193–200.

MacCartney, Bill; Manning, Christopher D. (2009): An Extended Model of Natural Logic. In: Proceedings of the Eighth International Conference on Computational Semantics (IWCS-8). Tilburg, The Netherlands, January 7-9, pp. 140–156.

MacGregor, Phil (2013): International News Agencies. Global Eyes that Never Blink. In *Journalism: New Challenges. Centre for Journalism & Communication Research, Bournemouth University.(PDF)*.

Magliano, Joseph P.; Graesser, Arthur C. (1991): A Three-pronged Method for Studying Inference Generation in Literary Text. In *Poetics* 20 (3), pp. 193–232.

Magnini, Bernardo; Zanoli, Roberto; Dagan, Ido; Eichler, Kathrin; Neumann, Guenter; Noh, Tae-Gil et al. (2014): The Excitement Open Platform for Textual Inferences. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations (ACL 2014). Baltimore, MD, USA, June 1. Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 43–48.

Mahdisoltani, Farzaneh; Biega, Joanna Asia; Suchanek, Fabian M. (2015): YAGO3: A Knowledge Base from Multilingual Wikipedias. In: Proceedings of the 7th Biennial Conference on Innovative Data Systems Research (CIDR 2015). Asilomar, CA, USA, January 4-7.

Maier, Scott R. (2005): Accuracy Matters. A Cross-Market Assessment of Newspaper Error and Credibility. In *Journalism & Mass Communication Quarterly* 82 (3), pp. 533–551.

Malakasiotis, Prodromos (2009): AUEB at TAC 2009. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Malakasiotis, Prodromos; Androutsopoulos (2007): Learning Textual Entailment Using SVMs and String Similarity Measures. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29, pp. 42–47.

Mani, Inderjeet; Verhagen, Marc; Wellner, Ben; Lee, Chong Min; Pustejovsky, James (2006): Machine Learning of Temporal Relations. In: Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the ACL (ACL 2006). Sydney, Australia, July 17-21. Morristown, NJ, USA: Association for Computational Linguistics, pp. 753–760.

Mani, Inderjeet; Wilson, George (2000): Robust Temporal Processing of News. In: Proceedings of the 38th Annual Meeting on Association for Computational Linguistics (ACL 2000). Hong Kong, China, October 3-6. Morristown, NJ, USA: Association for Computational Linguistics, pp. 69–76.

Mann, William C.; Thompson, Sandra A. (1988): Rhetorical Structure Theory. Toward a Functional Theory of Text Organization. In *Text - Interdisciplinary Journal for the Study of Discourse* 8 (3).

Manning, Christopher D. (2011): Part-of-Speech Tagging from 97% to 100%. Is It Time for Some Linguistics? In : Computational Linguistics and Intelligent Text Processing. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 171–189.

Manoff, Robert Karl; Schudson, Michael (Eds.) (1987): Reading the News. New York, NY: Pantheon Books.

Marcus, Mitchell; Marcinkiewicz, Mary Ann; Santorini, Beatrice (1993): Building a Large Annotated Corpus of English. The Penn Treebank. In *Computational Linguistics* 19 (2), pp. 314–330.

Mårdh, Ingrid (1980): Headlinese. On the Grammar of English Front Page Headlines. Zugl.: Lund, Diss. : 1980. Lund: Gleerup.

Markert, Katja; Nissim, Malvina (2005): Comparing Knowledge Sources for Nominal Anaphora Resolution. In *Computational Linguistics* 31 (3), pp. 367–402.

Markman, Ellen M. (1979): Realizing That You Don't Understand. Elementary School Children's Awareness of Inconsistencies. In *Child Development* 50 (3), p. 643.

Markman, Ellen M.; Gorin, Linda (1981): Children's Ability to Adjust their Standards for Evaluating Comprehension. In *Journal of Educational Psychology* 73 (3), pp. 320–325.

Marneffe, Marie-Catherine de; MacCartney, Bill; Manning, Christoper (2006): Generating Typed Dependency Parses from Phrase Structure Parses. In: Proceedings of International Conference on Language Resources and Evaluation (LREC 2006). Genoa, Italy, May 22-28.

Marneffe, Marie-Catherine de; Rafferty, Anna; Manning, Christoper (2008): Finding Contradictions in Text.

Màrquez, Lluís; Villarejo, Luis; Martí, M. A.; Taulé, Mariona (2007): SemEval-2007 Task 09: Multilevel Semantic Annotation of Catalan and Spanish. In: Proceedings of the 4th International Workshop on Semantic Evaluation (SemEval 2007). Prague, Czech Republic, June 23-24, pp. 42–47.

Marsic, Georgiana (2011): Temporal Processing of News: Annotation of Temporal Expressions, Verbal Events and Temporal Relations. Ph.D. Thesis. University of Wolverhampton, UK.

Massey, Brian L.; Levy, Mark R. (2016): Interactivity, Online Journalism, and English-Language Web Newspapers in Asia. In *Journalism & Mass Communication Quarterly* 76 (1), pp. 138–151.

Matveeva, Irina (2008): Generalized Latent Semantic Analysis for Document Representation. Ph.D. Thesis. Chicago, IL, USA: University of Chicago.

Mazur, Pavel; Dale, Robert (2008): A Rule Based Approach to Temporal Expression Tagging. In *Systems Science* 34 (4).

McCarthy, Joseph F.; Lehnert, Wendy G. (1995): Using Decision Trees for Conference Resolution. In: Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI 1995). Montreal, Canada, August 20-25, pp. 1050–1055.

McCready, Eric (2005): The Dynamics of Particles. Ph.D. Thesis: University of Austin.

McDowell, Ian (1992): Reuters Handbook for Journalists.

Medlock, Ben; Briscoe, Ted (2007): Weakly Supervised Learning for Hedge Classification in Scientific Literature. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics (ACL 2007). Prague, Czech Republic, June 25-27: Association for Computational Linguistics, pp. 992–999.

Mehdad, Yashar (2009): Automatic Cost Estimation for Tree Edit Distance Using Particle Swarm Optimization. In: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers (ACLShort 2009). Suntec, Singapore, August 4, pp. 289–292.

Mehdad, Yashar; Magnini, Bernardo (2009): Optimizing Textual Entailment Recognition Using Particle Swarm Optimization. In: Proceedings of the 2009 Workshop on Applied Textual Inference (TextInfer 2009). Suntec, Singapore, August 6, pp. 36–43.

Mel'čuk, Igor' A.; Beck, David; Polguère, Alain (2012): Semantics. From Meaning to Text. Amsterdam: John Benjamins Publishing Company.

Melucci, Massimo; Orio, Nicola (2003): A Novel Method for Stemmer Generation Based on Hidden Markov Models. In: Proceedings of the Twelfth ACM International Conference on Information and Knowledge Management (CIKM 2003). New Orleans, LA, USA, November 3-8. New York, New York, USA: ACM Press, pp. 131–138.

Mendelson, Elliott (1964): Introduction to Mathematical Logic. 3rd ed. New York usw: Van Nostrand Reinhold Company.

Mendes, Pablo N.; Jakob, Max; García-Silva, Andrés; Bizer, Christian (2011): DBpedia Spotlight. In: Proceedings of the 7th International Conference on Semantic Systems. Graz, Austria, September 7-9. New York, NY: ACM (ACM Digital Library), pp. 1–8.

Meyer, Josef; Dale, Robert (2002): Mining a Corpus to Support Associative Anaphora Resolution. In: Proceedings of the Fourth Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2002). Lisbon, Portugal, September 18-20.

Meyers, Adam; Reeves, Ruth; Macleod, Catherine; Szekely, Rachel; Zielinska, Veronika; Young, Brian; Grishman, Ralph (2004): The NomBank Project. An Interim Report. In: Proceedings of the NAACL-HLT 2004 Workshop: Frontiers in Corrpus Annotation (NAACL-HLT 2004). Boston, MA, USA, May 2-7, pp. 24–31.

Miestamo, Matti (2005c): Standard nNgation. The Negation of Declarative Verbal Main Clauses in a Typological Perspective. Berlin, New York: Mouton de Gruyter.

Miestamo, Matti (2006): On the Complexity of Standard Negation. In Mickael Suominen (Ed.): A Man of Measure. Festschrift in Honour of Fred Karlsson on His 60th Birthday. Helsinki, Finland: Linguistic Association of Finland, pp. 345–356.

Miestamo, Matti (2005a) (2005): Subtypes of Asymmetric Standard Negation. In Martin Haspelmath, Matthew S. Dryer, David Gil, Bernard Comrie (Eds.): World Atlas of Language Structures. Oxford: Oxford University Press, pp. 462–465.

Miestamo, Matti (2005b) (2005): Symmetric and Asymmetric Standard Negation. In Martin Haspelmath, Matthew S. Dryer, David Gil, Bernard Comrie (Eds.): World Atlas of Language Structures. Oxford: Oxford University Press, pp. 458–461.

Mihalcea, Rada; Corley, Courtney; Strapparava, Carlo (2006): Corpus-based and Knowledge-based Measures of Text Semantic Similarity. In: Proceedings of the Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference (AAAI 2006). Boston, MA, USA, July 16-20, pp. 775–780.

Mikheev, Andrei; Grover, Claire; Moens, Marc (1998): Description of the LTG System Used for MUC-7. In: Proceedings of the Seventh Message Understanding Conference (MUC-7). Fairfax, VA, USA, April 29-May 1.

Mikolov, Tomas; Sutskever, Ilya; Chen, Kai; Corrado, Greg; Dean, Jeffrey (2013): Distributed Representations of Words and Phrases and Their Compositionality. In: Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS 2013). Volume 2. Lake Tahoe, NV, USA, December 5-10, pp. 3111–3119.

Miller, George A.; Beckwith, Richard; Fellbaum, Christiane; Gross, Derek; Miller, Katherine J. (1990): Introduction to WordNet. An On-line Lexical Database. In *International Journal of Lexicography* 3 (4), pp. 235–244.

Miller, Lisa C. (1998): Power Journalism. Computer-assisted Reporting. Fort Worth: Harcourt Brace College Publishers.

Minard, Anne-Lyse; Speranza, Manuela; Sprugnoli, Rachele; Caselli, Tommaso (2015): FacTA. Evaluation of Event Factuality and Temporal Anchoring. In Cristina Bosco, Sara Tonelli, Fabio Massimo Zanzotto (Eds.): Proceedings of the Second Italian Conference on Computational Linguistics (CLiC-it 2015). Trento, Italy, December 03-04, pp. 187–192.

Minsky, Marvin L. (1975): A Framework for Representing Knowledge. In Patric Henry Winston (Ed.): The Psychology of Computer Vision. New York: McGraw-Hill, pp. 211–227.

Mitchell, Erika (2006): The Morpho-syntax of Negation and the Positions of NegP in the Finno-Ugric Languages. In *Lingua* 116 (3), pp. 228–244.

Mitkov, Ruslan (1998): Robust Pronoun Resolution with Limited Knowledge. In: Proceedings of the 17th International Conference on Computational Linguistics (COLING 1998). Montreal, Canada, August 10-14. Morristown, NJ, USA: Association for Computational Linguistics, pp. 869–875.

Mitkov, Ruslan (2002): Anaphora Resolution. London: Longman.

Mnih, Andriy; Hinton, Geoffrey E. (2009): A Scalable Hierarchical Distributed Language Model. In: Proceedings of the Conference Advances in Neural Information Processing Systems 21 (NIPS 2008). Vancouver, Canada, December 8-11: Curran Associates, Inc, pp. 1081–1088.

Mohammad, Saif; Dorr, Bonnie J.; Egan, Melissa; Madnani, Nitin; Zajic, David; Lin, Jimmy (2008): Multiple Alternative Sentence Compressions and Word-Pair Antonymy for Automatic Text Summarization and Recognizing Textual Entailment. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Montague, Richard; Thomason, Richmond H. (1974): Formal philosophy. Selected Papers of Richard Montague. New Haven, London: Yale University Press.

Montalvo-Huhn, Orlando; Taylor, Stephen (2008): Textual Entailment – Fitchburg State College. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Mooney, Raymond J. (2004): Machine Learning. In Ruslan Mitkov (Ed.): The Oxford Handbook of Computational Linguistics. [Pbk. ed.]. Oxford, New York: Oxford University Press, pp. 376–394.

Moore, George. E.; Baldwin, Thomas (Eds.) (1993): G. E. Moore. Selected Writings. London: Routledge.

Morante, Roser; Daelemans, Walter (2009b): A Metalearning Approach to Processing the Scope of Negation. In: Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL 2009). Boulder, CO, USA, June 4-5. Stroudsburg, Pa.: Association for Computational Linguistics, p. 21.

Morante, Roser; Daelemans, Walter (2009a): Learning the Scope of Hedge Cues in Biomedical Texts. In: Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing (BioNLP 2009). Boulder, CO, USA, June 4-5. East Shroudburg, PA, USA: Association of Computational Linguistics, pp. 28–36.

Morante, Roser; Daelemans, Walter (2012): ConanDoyle-neg: Annotation of Negation in Conan Doyle stories. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012). Istanbul, Turkey, May.

Morante, Roser; Sporleder, Caroline (2012): Modality and Negation. An Introduction to the Special Issue. In *Computational Linguistics* 38 (2), pp. 223–260.

Morris, Jane; Hirst, Graeme (1991): Lexical Cohesion Computed by Thesaural Relations as an Indicator of the Structure of Text. In *Computational Linguistics* 17 (1), pp. 21–48.

Morsey, Mohamed; Lehmann, Jens; Auer, Sören; Stadler, Claus; Hellmann, Sebastian (2012): DBpedia and the Live Extraction of Structured Data from Wikipedia. In *Program: electronic library and information systems* 46 (2), pp. 157–181.

Mučnik, Beniamin S. (1985): Čelovek i tekst. Osnovy kul'tury pis'mennoj reči. Moskva: Kniga.

Mutalik, P. G.; Deshpande, A.; Nadkarni, P. M. (2001): Use of General-purpose Negation Detection to Augment Concept Indexing of Medical Documents. A Quantitative Study Using the UMLS. In *Journal of the American Medical Informatics Association* 8 (6), pp. 598–609.

Nairn, Rowan; Condoravdi, Cleo; Karttunen, Lauri (2006): Computing Relative Polarity for Textual Inference.

Neverilova, Zuzana (2014): Paraphrase and Textual Entailment Generation in Czech. In *Computación y Sistemas* 18 (3).

Ng, Hwee Tou; Zhou, Yu; Dale, Robert; Gardiner, Mary (2005): A Machine Learning Approach to Identification and Resolution of One-Anaphora. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI 2005). Edinburgh, Scotland, July 30-August 5, pp. 1105–1110.

Ng, Vincent; Cardie, Claire (2002): Improving Machine Learning Approaches to Coreference Resolution. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL 2002). Philadelphia, PA, USA, July 7-12. Morristown, NJ, USA: Association for Computational Linguistics, p. 104.

Ng, Vincent; Cardie, Claire (2002a) (2002): Identifying Anaphoric and Non-anaphoric Noun Phrases to Improve Coreference Resolution. In: Proceedings of the 19th International Conference on Computational Linguistics (COLING 2002). Taipei, Taiwan, August 24-September 1. Morristown, NJ, USA: Association for Computational Linguistics, pp. 1–7.

Nicholas, David W.; Trabasso, Tom (1980): Towards a Taxonomy of Inferences. In Friedrich Wilkening (Ed.): Information Integration by Children. Hillsdale, N.J.: Erlbaum, pp. 243–266.

Nirenburg, Sergei; Raskin, Victor (2004): Ontological Semantics. Cambridge, Mass., London: MIT Press.

Nørgård-Sørensen, Jens (1992): Coherence Theory. The Case of Russian. Berlin, New York: Mouton de Gruyter.

Norrick, Neal R. (1986): A Frame-theoretical Analysis of Verbal Humor. Bisociation as Schema Conflict. In *ACM SIGIR Forum* 60 (3-4).

Ohler, Josef (2016): Die Nachrichtensprache. In Dietz Schwiesau, Josef Ohler (Eds.): Nachrichten - klassisch und multimedial. Wiesbaden: Springer Fachmedien Wiesbaden, 99-158.

Ohler, Josef; Schwiesau, Dietz (2016): Der Nachrichtenaufbau. In Dietz Schwiesau, Josef Ohler (Eds.): Nachrichten - klassisch und multimedial. Wiesbaden: Springer Fachmedien Wiesbaden, 33-57.

Oltramari, Alessandro; Gangemi, Aldo; Guarino, Nicola; Masolo, Claudio (2002): Restructuring WordNet's Top-Level: The OntoClean Approach. In: Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002). Las Palmas, Spain, May 29-31. Las Palmas, Canary Islands, Spain: European Language Resources Association (ELRA), pp. 17–26.

Opgenhaffen, Michael; d'Haenens, Leen (2011): The Impact of Online News Features on Learning from News: A Knowledge Experiment. In *International Journal of Internet Science* 6 (1), pp. 8–28.

Orwell, George (1961): Politics and the English Language (1946). Repr. In : Collected Essays. London: Heinemann.

O'Sullivan, John (2005): Ireland: Newsprint Colonizes the Web. In R. van der Wurff, Edmund Lauf (Eds.): Print and Online Newspapers in Europe. A Comparative Analysis in 16 Countries. Amsterdam, Somerset NJ: Het Spinhuis; Distribution in the USA Canada Latin America and the UK Transaction Publishers, pp. 159–172.

Otero, Jose C.; Campanario, Juan M. (1990): Comprehension Evaluation and Regulation in Learning from Science Texts. In *Journal of Research in Science Teaching* 27 (5), pp. 447–460.

Otero, José; Kintsch, Walter (1992): Failures to Detect Contradictions in a Text. What Readers Believe versus What They Read. In *Psychological Science* 3 (4), pp. 229–236.

Ovchinnikova, Ekaterina (2012): Integration of World Knowledge for Natural Language Understanding. 1. Aufl. s.l.: Springer-Verlag.

Özgür, Arzucan; Radev, Dragomir R. (2009): Detecting Speculations and Their Scopes in Scientific Text. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing Volume 3 (EMNLP 2009). Singapore, August 6-7. Morristown, NJ, USA: Association for Computational Linguistics, p. 1398.

Padó, Sebastian; Marneffe, Marie-Catherine de; MacCartney, Bill; Rafferty, Anna; Yeh, Eric; Manning, Christoper (2008): Deciding Entailment and Contradiction with Stochastic and Edit Distance-based Alignment. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Padó, Sebastian; Noh, Tae-Gil; Stern, Asher; Wang, Rui; Zanoli, Roberto (2015): Design and Realization of a Modular Architecture for Textual Entailment. In *Natural Language Engineering* 21 (2), pp. 167–200.

Paice, Chris D. (1990): Another Stemmer. In *ACM SIGIR Forum* 24 (3), pp. 56–61.

Palmer, Frank R. (1986): Mood and Modality. Cambridge: Cambridge University Press.

Palmer, Martha; Gildea, Daniel; Kingsbury, Paul (2005): The Proposition Bank. An Annotated Corpus of Semantic Roles. In *Computational Linguistics* 31 (1), pp. 71–106.

Pang, Bo; Lee, Lillian; Vaithyanathan, Shivakumar (2002): Thumbs up? Sentiment Classification Using Machine Learning Techniques. In: Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP 2002). Pennsylvania, PA, USA, July 6-7. Morristown, NJ, USA: Association for Computational Linguistics, pp. 79–86.

Paris, Scott G.; Myers, Meyer [II] (1981): Comprehension Monitoring, Memory, and Study Strategies of Good and Poor Readers. In *Journal of Reading Behavior* 13 (1), pp. 5–22.

Parry, William T.; Hacker, Edward A. (1991): Aristotelian Logic. Albany, NY: State University of New York Press.

Parsons, Terence (1990): True Contradictions. In *Canadian Journal of Philosophy* 20 (3), pp. 335–353.

Parsons, Terence (2017): The Traditional Square of Opposition. In Edward N. Zalta (Ed.): The Stanford Encyclopedia of Philosophy. Summer 2017: Metaphysics Research Lab, Stanford University.

Payne, John R. (1985): Negation. In Timothy Shopen (Ed.): Language Typology and Syntactic Description: Volume 1. Clause Structure. Cambridge: Cambridge University Press, pp. 197–242.

Penka, Doris (2011): Negative Indefinites. Zugl.: Tübingen, Univ., Diss., 2007. Oxford: Oxford University Press.

Pennington, Jeffrey; Socher, Richard; Manning, Christopher D. (2014): GloVe. Global Vectors for Word Representation. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2014). Doha, Qatar, October 25–29, pp. 1532–1543.

Perfetti, Charles A.; Stafura, Joseph Z. (2015): Comprehending Implicit Meanings in Text Without Making Inferences. In Edward J. O'Brien, Anne E. Cook, Robert F. Lorch (Eds.): Inferences during Reading. Cambridge: Cambridge University Press, pp. 1–18.

Perina, Alessandro; Jojic, Nebojsa; Bicego, Manuele; Truski, Andrzej (2013): Documents as Multiple Overlapping Windows into Grids of Counts. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Q. Weinberger (Eds.): Advances in Neural Information Processing Systems 26: Curran Associates, Inc, pp. 10–18.

Petrov, Slav; Barrett, Leon; Thibaux, Romain; Klein, Dan (2006): Learning Accurate, Compact, and Interpretable Tree Annotation. In: Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the ACL (ACL 2006). Sydney, Australia, July 17-21. Morristown, NJ, USA: Association for Computational Linguistics, pp. 433–440.

Pham, Minh Quang Nhat; Le Nguyen, Minh; Shimazu, Akira (2013): Using Shallow Semantic Parsing and Relation Extraction for Finding Contradiction in Text. In: Proceedings of the Sixth International Joint Conference on Natural Language Processing (IJCNLP 2013). Nagoya, Japan, October: Asian Federation of Natural Language Processing, pp. 1017–1021.

Poesio, Massimo; Kabadjov, Mijail A. (2004): A General-Purpose, Off-the-shelf Anaphora Resolution Module: Implementation and Preliminary Evaluation. In: Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004). Lisbon, Portugal, May 26-28, pp. 663–666.

Porter, Martin F. (1980): An Algorithm for Suffix Stripping. In *Program* 14 (3), pp. 130–137.

Potts, Christopher (2002): The Lexical Semantics of Parenthical-as and Appositive-which. In *Syntax* 5 (1), pp. 55–88.

Potts, Christopher (2002): The Syntax and Semantics of As-Parentheticals. In *Natural Language and Linguistic Theory* 20 (3), pp. 623–689.

Potts, Christopher (2015): Presupposition and Implicature. In Shalom Lappin, Chris Fox (Eds.): The Handbook of Contemporary Semantic Theory. Chichester, UK: John Wiley & Sons, Ltd, pp. 168–202.

Prabhakaran, Vinodkumar; Rambow, Owen; Diab, Mona (2010): Automatic Committed Belief Tagging. In: Proceedings of the 23th International Conference on Computational Linguistics (COLING 2010). Beijing, China, August 23-27, pp. 1014–1022.

Pradhan, Sameer; Loper, Edward; Dligach, Dmitriy; Palmer, Martha (2007): SemEval-2007 Task 17: English Lexical Sample, SRL and All Words. In: Proceedings of the 4th International Workshop on Semantic Evaluation (SemEval 2007). Prague, Czech Republic, June 23-24, pp. 87–92.

Pradhan, Sameer; Moschitti, Alessandro; Xue, Nianwen; Uryupina, Olga; Zhang, Yuchen (2012): CoNLL-2012 Shared Task: Modeling Multilingual Unrestricted Coreference in OntoNotes. In: Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2012). Jeju Island, Korea, July 12-14, pp. 1–40.

Priest, Graham (1998): What is so Bad about Contradictions? In *The Journal of Philosophy* 95 (8), p. 410.

Priest, Graham (1999): Perceiving Contradictions. In *Australasian Journal of Philosophy* 77 (4), pp. 439–446.

Prince, Ellen F. (1981): Toward a Taxonomy of Given–New Information. In Peter Cole (Ed.): Radical Pragmatics. New York: Academic Press, pp. 223–255.

Prior, Arthur Norman (1967): Negation. In Paul Edwards (Ed.): The Encyclopedia of Philosophy. New York: Macmillan.

Punyakanok, Vasin; Roth, Dan; Yih, Wen-tau (2008): The Importance of Syntactic Parsing and Inference in Semantic Role Labeling. In *Computational Linguistics* 34 (2), pp. 257–287.

Pustejovsky, James; Hanks, Patrick; Saurí, Roser; See, Andrew; Gaizauskas, Robert; Setzer, Andrea et al. (2003): The TIMEBANK Corpus. In: Proceedings of the Corpus Linguistics 2003 Conference (ACL 2003). Lancaster, UK, March 28-31. Lancaster: UCREL, 647-456.

Quillian, M. Ross (1968): Semantic Memory. In Marvin L. Minsky (Ed.): Semantic Information Processing. Cambridge, London: MIT Press, pp. 227–270.

Quine, Willard V. O. (1959): Methods of Logic. Rev. ed. New York: Holt Rinehart Winston.

Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey; Svartvik, Jan (2010): A Comprehensive Grammar of the English Language. 1. international ed. Harlow: Longman.

Raghunathan, Karthik; Lee, Heeyoung; Rangarajan, Sudarshan; Chambers, Nathanael; Surdeanu, Mihai; Jurafsky, Dan; Manning, Christopher (2010): A Multi-Pass Sieve for Coreference Resolution. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2010).

Raskin, Victor (1985): Semantic Mechanisms of Humor. Dordrecht: Reidel.

Reah, Danuta (2003): The Language of Newspapers. 2. ed., Repr. London: Routledge.

Recasens, Marta; Marneffe, Marie-Catherine de; Potts, Christopher (2013): The Life and Death of Discourse Entities. Identifying Singleton Mentions. In: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2013). Atlanta, GA, USA, June 9-15, pp. 627–633.

Reddick, Randy; King, Elliot (2001): The Online Journalist. Using the Internet and Other Electronic Resources. 3rd ed. Fort Worth: Harcourt College Publishers.

Regenbogen, Arnim; Meyer, Uwe (1998): Wörterbuch der philosophischen Begriffe. Hamburg: F. Meiner Verlag.

Reichenbach, Hans (1947): Elements of Symbolic Logic. New York: Free Press.

Reinhard, Tanya (1976): Polarity Reversal: Logic or Pragmatics? In *Linguistic Inquiry* 7, pp. 697–705.

Reis, Ron; Spekman, Nancy J. (1983): The Detection of Reader-Based versus Text-Based Inconsistencies and the Effects of Direct Training of Comprehension Monitoring among Upper-Grade Poor Comprehenders. In *Journal of Reading Behavior* 15 (2), pp. 49–60.

Reumann, Kurt (2002): Journalistische Darstellungsformen. In Elisabeth Noelle-Neumann, Winfried Schulz, Jürgen Wilke (Eds.): Das Fischer-Lexikon. Publizistik Massenkommunikation. Frankfurt am Main: Fischer.

Reynar, Jeffrey C.; Ratnaparkhi, Adwait (1997): A Maximum Entropy Approach to Identifying Sentence Boundaries. In: Proceedings of the fifth conference on Applied Natural Language Processing (ANLC 1997). Washington, DC, March 31-April 3. Morristown, NJ, USA: Association for Computational Linguistics, pp. 16–19.

Richardson, John E. (2007): Analysing Newspapers. An Approach from Critical Discourse Analysis. Houndmills, New York: Palgrave Macmillan.

Richardson, Stephen D.; Dolan, William B.; Vanderwende, Lucy (1998): MindNet: Acquiring and Structuring Semantic Information from Text. In: Proceedings of the 17th International Conference on Computational Linguistics (COLING 1998). Montreal, Canada, August 10-14. Morristown, NJ, USA: Association for Computational Linguistics, pp. 1098–1102.

Rieger, Reinhold (2005): Contradictio. Theorien und Bewertungen des Widerspruchs in der Theologie des Mittelalters. Zugl.: Tübingen, Univ., Habil.-Schr., 2004. Tübingen: Mohr Siebeck.

Rieger, Charles J. (1975): Conceptual Memory and Inference. In Roger C. Schank (Ed.): Conceptual Information Processing: North-Holland, pp. 157–288.

Riemer, Nick (2010): Introducing Semantics. Cambridge: Cambridge University Press.

Ripley, David (2011): Contradictions at the Borders. In Rick Nouwen, Robert van Rooij, Uli Sauerland, Hans-Christian Schmitz (Eds.): Vagueness in Communication. International Workshop, ViC 2009, held as part of ESSLLI 2009, Bordeaux, France, July 20 - 24, 2009 ; revised selected papers. Berlin: Springer, pp. 169–188.

Ripley, David (2013): Sorting out the Sorites. In Koji Tanaka, Francesco Berto, Edwin Mares, Francesco Paoli (Eds.): Paraconsistency: Logic and Applications. Dordrecht: Springer Netherlands, pp. 329–348.

Ritter, Alan; Downey, Doug; Soderland, Stephen; Etzioni, Oren (2008): It's a Contradiction - No, it's not: A Case Study Using Functional Relations. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2008). Honolulu, HI, USA, October 25-27, pp. 11–20.

Roberts, Craige (2003): Uniqueness Presuppositions in English Definite Noun Phrases. In *Linguistics and Philosophy* 26 (3), pp. 287–350.

Roberts, Craige (2004): Context in Dynamic Interpretation. In Laurence R. Horn, Gregory L. Ward (Eds.): The Handbook of Pragmatics. Malden MA: Blackwell Pub, pp. 197–220.

Robertson, Stephen; Walker, Steve; Jones, Susan; Hancock-Beaulieu, Micheline; Gatford, Mike (1995): Okapi at Trec-3. In: Overview of the Third Text REtrieval Conference (TREC-3). Gaithersburg, MD, USA, November 02-04, pp. 109–126.

Rochemont, Michael S.; Halvorsen, Per-Kristian (1980): The Syntax and Semantics of Cleft Constructions. In *Language* 56 (3), p. 694.

Roeh, Itzhak (1982): The Rhetoric of News in the Israel Radio. Some Implications of Language and Style for Newstelling; [a Case Study with Implications for Western Newstelling]. Bochum: Studienverlag Brockmeyer.

Roeh, Itzhak; Feldman, Saul (1984): The Rhetoric of Numbers in Front-page Journalism. How Numbers Contribute to the Melodramatic in the Popular Press. In *Text - Interdisciplinary Journal for the Study of Discourse* 4 (4).

Roshco, Bernard (1975): Newsmaking. Chicago: University of Chicago Press.

Routley, Richard; Routley, Val (1985): Negation and Contradiction. In *Revista Colombiana de Matemáticas* 19 (1-2), pp. 201–231.

Rubenstein, Herbert; Goodenough, John B. (1965): Contextual Correlates of Synonymy. In *Communications of the ACM* 8 (10), pp. 627–633.

Rubin, Victoria L. (2007): Stating with Certainty or Stating with Doubt. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2007). Rochester, New York. Morristown, NJ, USA: Association for Computational Linguistics, pp. 141–144.

Rubin, Victoria L.; Liddy, Elizabeth D.; Kando, Noriko (2005): Certainty Identification in Texts. Categorization Model and Manual Tagging Results. In James G. Shanahan, Yan Qu, Janyce Wiebe (Eds.): Computing Attitude and Affect in Text: Theory and Applications, vol. 20. Berlin/Heidelberg: Springer Science & Business Media, pp. 61–76.

Rubner, Yossi; Tomasi, Carlo; Guibas, Leonidas J. (1998): A Metric for Distributions with Applications to Image Databases. In: Proceedings of the Sixth International Conference on Computer Vision (ICCV 1998). Bombay, India, January 04-07, 59-67.

Ruppenhofer, Josef; Sunde, Jonas; Pinkal, Manfred (2010): Generating FrameNets of Various Granularities: The FrameNet Transformer. In: Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC 2010). Valletta, Malta.

Russell, Bertrand (1905): On Denoting. In *Mind* 14 (4), pp. 479–493.

Russell, Bertrand (1957): Mr. Strawson on Referring. In *Mind* 66, pp. 385–389.

Sainsbury, Richard M. (1991): Logical Forms. An Introduction to Philosophical Logic. Oxford: Blackwell.

Salas-Zárate, María Del Pilar; Medina-Moreira, José; Lagos-Ortiz, Katty; Luna-Aveiga, Harry; Rodríguez-García, Miguel Ángel; Valencia-García, Rafael (2017): Sentiment Analysis on Tweets about Diabetes. An Aspect-Level Approach. In *Computational and mathematical methods in medicine* 2017.

Salton, Gerard; Buckley, Christopher (1988): Term-weighting Approaches in Automatic text Retrieval. In Information Processing and Management 24 (5), pp. 513–523.

Sanders, Ted J. M.; Spooren, Wilbert P. M.; Noordman, Leo G. M. (1992): Toward a Taxonomy of Coherence Relations. In *Discourse Processes* 15 (1), pp. 1–35.

Šarić, Frane; Glavaš, Goran; Karan, Mladen; Šnajder, Jan; Dalbelo Bašić, Bojana (2012): TakeLab: Systems for Measuring Semantic Text Similarity. In: Proceedings of the First Joint Conference on Lexical and Computational Semantics (SEM 2012). Montreal, Canada, June 7-8, pp. 441–448.

Sauerland, Uli (2011): Vagueness in Language: The Case Against Fuzzy Logic Revisited. In Petr Cintula, Christian G. Fermueller, Lluis Godo (Eds.): Understanding Vagueness. Logical, Philosophical, and Linguistic Perspectives. London: College Publications, pp. 185–198.

Sauerland, Uli; Schenner, Mathias (2007): Shifting Evidentials in Bulgarian. In *Sinn und Bedeutung* 11, pp. 525–539.

Saurí, Roser; Knippen, Robert; Verhagen, Marc; Pustejovsky, James (2005): Evita: A Robust Event Recognizer for QA Systems. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT-EMNLP 2005). Vancouver, Canada, October 6-8. Morristown, NJ, USA: Association for Computational Linguistics, pp. 700–707.

Saurí, Roser; Pustejovsky, James (2007): Determining Modality and Factuality for Text Entailment. In: Proceedings of the International Conference on Semantic Computing (ICSC 2007). Irvine, CA, USA, September 17-19: IEEE, pp. 509–516.

Saurí, Roser; Pustejovsky, James (2009): FactBank. A Corpus Annotated with Event Factuality. In *Lang Resources & Evaluation* 43 (3), pp. 227–268.

Saurí, Roser; Pustejovsky, James (2012): Are You Sure That This Happened? Assessing the Factuality Degree of Events in Text. In *Computational Linguistics* 38 (2), pp. 261–299.

Schank, Roger C. (1972): Conceptual Dependency. A Theory of Natural Language Understanding. In *Cognitive Psychology* 3 (4), pp. 552–631.

Schank, Roger C. (1975): Using Knowledge to Understand. In: Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing (TINLAP 1975). Cambridge, Massachusetts. Morristown, NJ, USA: Association for Computational Linguistics, pp. 117–121.

Schank, Roger C.; Abelson, Robert P. (1977): Scripts, Plans, Goals and Understanding. An Inquiry into Human Knowledge Structures. Hillsdale, NJ: Erlbaum.

Schiller, Dan (1981): Objectivity and the News. The Public and the Rise of Commercial Journalism. Philadelphia: Pennsylvania University Press.

Schlenker, Philippe (2009): Local Contexts. In *Semantics and Pragmatics* 2.

Schmid, Helmut (2010): Decision Trees. In Alexander Clark, Chris Fox, Shalom Lappin (Eds.): The Handbook of Computational Linguistics and Natural Language Processing. Oxford, UK: Wiley-Blackwell, pp. 180–196.

Schmolz, Helene (2013): Anaphora Resolution and Text Retrieval. Dissertation. Available online at http://dx.doi.org/10.1515/9783110416756.

Schneider, Wolf; Raue, Paul-Josef (2009): Das neue Handbuch des Journalismus. 5. Aufl. Reinbek: Rowohlt Taschenbuch-Verlag.

Schoubye, Anders J. (2009): Descriptions, Truth Value Intuitions, and Questions. In *Linguistics and Philosophy* 32 (6), pp. 583–617.

Schröder, Thomas (2003): Die Handlungsstruktur von Texten. Ein integrativer Beitrag zur Texttheorie. Zugl.: Tübingen, Univ., Habil.-Schr., 2000. Tübingen: Narr.

Schubert, Lenhart (2015): Computational Linguistics. In Edward N. Zalta (Ed.): The Stanford Encyclopedia of Philosophy. Spring 2015: Metaphysics Research Lab, Stanford University.

Schudson, Michael (1995): Discovering the News. A Social History of American Newspapers. [Nachdr.]. New York, NY: Basic Books.

Schulz, Moritz (2010): Epistemic Modals and Informational Consequence. In *Synthese* 174 (3), pp. 385–395.

Schulz, Thomas R. (1976): A Cognitive-developmental Analysis of Humour. In Antony J. Chapman (Ed.): Humour and Laughter. Theory, Research and Applications. London: Wiley, pp. 11–36.

Schwarz, Florian (2009): Two Types of Definites in Natural Language. Ph.D. Thesis: University of Massachussets Amherst.

Schwiesau, Dietz; Ohler, Josef (2013): Die Nachricht. In Presse, Radio, Fernsehen, Nachrichtenagentur und Internet: Ein Handbuch für Ausbildung und Praxis. Wiesbaden: Springer Fachmedien.

Schwiesau, Dietz; Ohler, Josef (Eds.) (2016): Nachrichten - klassisch und multimedial. Wiesbaden: Springer Fachmedien Wiesbaden.

Scott, William A. (1955): Reliability of Content Analysis. The Case of Nominal Scale Coding. In *Public Opinion Quarterly* 19 (3), pp. 127–141.

Sekine, Satoshi; Sudo, Kiyoshi; Nobata, Chikashi (2002): Extended Named Entity Hierarchy. In: Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002). Las Palmas, Spain, May 29-31. Las Palmas, Canary Islands, Spain: European Language Resources Association (ELRA).

Setzer, Andrea (2001): Temporal Information in Newswire Articles: an Annotation. Scheme and Corpus Study. Ph.D. Thesis. Sheffield University.

Sherman, Mark A. (1973): Bound to be Easier? The Negative Prefix and Sentence Comprehension. In *Journal of Verbal Learning and Verbal Behavior* 12 (1), pp. 76–84.

Shih, Chengwei; Lee, Chengwei; Tsai, Richard Tzonghan; Hsu, Wenlian (2012): Validating Contradiction in Texts Using Online Co-Mention Pattern Checking. In *ACM Transactions on Asian Language Information Processing* 11 (4), pp. 1–21.

Shnarch, Eyal; Barak, Libby; Dagan, Ido (2009): Extracting Lexical Reference Rules from Wikipedia. In : Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and 4th International Joint Conference on Natural Language Processing of the AFNLP (ACL-IJCNLP 2009). Suntec, SIngapore, August 2-7. Singapore, pp. 450–458.

Siblini, Reda; Kosseim, Leila (2008): Using Ontology Alignment for the TAC RTE Challenge. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Siegel, Eric V.; McKeown, Kathleen R. (2001): Learning Methods to Combine Linguistic Indicators: Improving Aspectual Classification and Revealing Linguistic Insights. In *Computational Linguistics* 26 (4), pp. 595–628.

Sigal, Leon V. (1987): Sources Make the News. In Robert Karl Manoff, Michael Schudson (Eds.): Reading the News. New York, NY: Pantheon Books, pp. 9–37.

Silverman, Craig (2007): Regret the Rrror. How Media Mistakes Pollute the Press and Imperil Free Speech. New York: Union Square Press.

Silvia, Tony (Ed.) (2001): Global News. Perspectives on the Information Age. 2nd ed. Ames, Iowa: Iowa State University Press.

Simons, Mandy (2001): On the Conversational Basis of Some Presuppositions. In *SALT* 11, p. 431.

Simons, Mandy (2003): Presupposition and Accommodation: Understanding the Stalnakerian Picture. In *Philosophical Studies* 112 (3), pp. 251–278.

Singer, Jane B. (2016): Who are These Guys? In *Journalism: Theory, Practice & Criticism* 4 (2), pp. 139–163.

Singer, Murray (1988): Inferences in Reading. In Gary Waller, G. E. MacKinnon, Meredyth Danemann (Eds.): Reading Research. Advances in Theory and Practice. New York, N.Y.: Academic Press, pp. 177–219.

Singer, Murray (1996): Comprehending Consistent and Inconsistent Causal Text Sequences. A Construction-Integration Analysis. In *Discourse Processes* 21 (1), pp. 1–21.

Singer, Murray; Graesser, Arthur C.; Trabasso, Tom (1994): Minimal or Global Inference during Reading. In *Journal of Memory and Language* 33 (4), pp. 421–441.

Smith, Nicholas Jeremy Josef (2008): Vagueness and Degrees of Truth. Oxford: Oxford University Press.

Smith, Tony C.; Frank, Eibe (2016): Introducing Machine Learning Concepts with WEKA. In *Methods in Molecular Biology* 1418, pp. 353–378.

Smyth, Herbert Weir (1974): Greek Grammar. 9. ed. Cambridge, Mass.: Harvard University Press.
Snider, Todd (2014): Using Tautologies and Contradictions. In: Proceedings of the 19th Sinn und Bedeutung Conference (SuB 2015). Göttingen, Germany, September, pp. 590–607.
Soames, Scott (1982): How Presuppositions are Inherited: A Solution to the Projection Problem. In *Linguistic Inquiry* 13 (4), pp. 483–545.
Socher, Richard; Perelygin, Alex; Wu, Jean; Chuang, Jason; Manning, Christopher D.; Ng, Andrew; Potts, Christopher (2013): Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2013). Seattle, USA, October 18-21.
Soon, Wee Meng; Ng, Hwee Tou; Lim, Daniel Chung Yong (2001): A Machine Learning Approach to Coreference Resolution of Noun Phrases. In *Computational Linguistics* 27 (4), pp. 521–544.
Sowa, John F. (1987): Semantic Networks. In Stuart C. Shapiro, David Eckroth (Eds.): Encyclopedia of Artificial Intelligence. New York, Chichester: Wiley.
Speer, Robert; Havasi, Catherine (2012): Representing General Relational Knowledge in ConceptNet 5. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012). Istanbul, Turkey, May.
Stalnaker, Robert (1973): Presuppositions. In *Journal of Philosophical Logic* 2 (4).
Stalnaker, Robert (1998): On the Representation of Context. In *Journal of Logic, Language and Information* 7 (1), pp. 3–19.
Stalnaker, Robert C. (1970): Pragmatics. In *Synthese* 22 (1-2), pp. 272–289.
Stalnaker, Robert C. (1972): Pragmatics. In Donald Davidson, Gilbert Harman (Eds.): Semantics of Natural Language. 2. ed., 3. pr. Dordrecht: Reidel, pp. 380–397.
Stalnaker, Robert C. (1974): Pragmatic Presuppositions. In Milton Karl Munitz, Peter K. Unger (Eds.): Semantics and Philosophy. New York: New York University Press, pp. 141–177.
Stalnaker, Robert C. (1978): Assertion. In Peter Cole (Ed.): Pragmatics. New York: Academic Press, pp. 315–332.
Stede, Manfred (2007): Korpusgestützte Textanalyse. Grundzüge der Ebenen-orientierten Textlinguistik. Tübingen: Narr.
Stenvall, Maija (2008): Unnamed Sources as Rhetorical Constructs in News Agency Reports. In *Journalism Studies* 9 (2), pp. 229–243.
Stern, Asher; Dagan, Ido (2012): BiuTee: a Modular Open-source System for Recognizing Textual Entailment. In: Proceedings of the ACL 2012 System Demonstrations (ACL 2012). Jeju Island, Korea, July 10, pp. 73–78.
Stoyanov, Veselin; Cardie, Claire; Gilbert, Nathan; Riloff, Ellen; Buttler, David; Hysom, David (2010): Coreference Resolution with Reconcile. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics Conference Short Papers (ACLShort 2010). Uppsala, Sweden, July 11-16, pp. 156–161.
Strawson, Peter F. (1950): On Referring. In *Mind* 59, pp. 320–344.
Strawson, Peter F. (1952): Introduction to Logical Theory. [Nachdr. d. Ausg.] Methuen 1952. London: Methuen.
Strötgen, Jannik; Gertz, Michael (2010): HeidelTime: High Quality Rule-based Extraction and Normalization of Temporal Expressions. In: Proceedings of the 5th International Workshop on Semantic Evaluation (SemEval 2010). Los Angeles, CA, USA, July 15-16, pp. 321–324.
Strube, Michael (2010): Anaphernresolution. In Kai-Uwe Carstensen, Christian Ebert, Cornelia Ebert, Susanne Jekat, Ralf Klabunde, Hagen Langer (Eds.): Computerlinguistik und Sprachtechnologie. Eine Einführung. 3., überarb. und erw. Aufl. Heidelberg: Spektrum Akad. Verlag, pp. 399–409.
Stuckardt, Roland (2001): Design and Enhanced Evaluation of a Robust Anaphor Resolution Algorithm. In *Computational Linguistics* 27 (4), pp. 479–506.
Suchanek, Fabian M.; Kasneci, Gjergji; Weikum, Gerhard (2007): Yago: A Core of Semantic Knowledge. In: Proceedings of the 16th International Conference on World Wide Web (WWW 2007). Banff, Canada, May 8-12. New York, NY: ACM (ACM Digital Library), pp. 697–706.
Sultan, Md Arafat; Bethard, Steven; Sumner, Tamara (2015): DLS@CU. Sentence Similarity from Word Alignment and Semantic Vector Composition. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). Denver, CO, USA, June 4-5.
Sun, Yufei; Ma, Liangli; Wang, Shuang (2015): A Comparative Evaluation of String Similarity Metrics for Ontology Alignment. In Journal of Information and Computational Science 12 (3), pp. 957–964.
Svintsov, Vitali Ivanovich (1979): Smislovoi analis i obrabotka teksta: ot rukopisi k knige. Moskva: Kniga.
Swart, Henriëtte de (2010): Expression and Interpretation of Negation. An OT Typology. Dordrecht: Springer Science+Business Media B.V.

Swovelin, Barbara V. (2006): CliffsAP: English Language and Composition. 3rd ed. Hoboken, N.J., Chichester: Wiley.

Szarvas, György (2008): Hedge Classification in Biomedical Texts with a Weakly Supervised Selection of Keywords. In: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL 2008). Columbus, OH, USA, June 15-20, pp. 281–289.

Taboada, Maite; Mann, William (2006): Rhetorical Structure Theory. Looking Back and Moving Ahead. In *Discourse Studies* 8 (3), pp. 423–459.

Takabatake, Yu; Morita, Hajime; Kawahara, Daisuke; Kurohashi, Sadao; Higashinaka, Ryuichiro; Matsuo, Yoshihiro (2015): Classification and Acquisition of Contradictory Event Pairs using Crowdsourcing. In: Proceedings of the The 3rd Workshop on EVENTS: Definition, Detection, Coreference, and Representation. Denver, CO, USA, June 1. Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 99–107.

Tatu, Marta; Moldovan, Dan (2006): A Logic-based Semantic Approach to Recognizing Textual Entailment. In: Proceedings of the COLING/ACL on Main Conference Poster Sessions (COLING-ACL 2006). Sydney, Australia, July 17-18, pp. 819–826.

Tatu, Marta; Moldovan, Dan (2007): COGEX at RTE3. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29, pp. 22–27.

Tausczik, Yla R.; Pennebaker, James W. (2010): The Psychological Meaning of Words. LIWC and Computerized Text Analysis Methods. In *Journal of Language and Social Psychology* 29 (1), pp. 24–54.

Tesnière, Lucien (2015): Elements of Structural Syntax. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Tetreault, Joel R. (2001): A Corpus-based Evaluation of Centering and Pronoun Resolution. In *Computational Linguistics* 27 (4), pp. 507–520.

Thelwall, Mike; Buckley, Kevan; Paltoglou, Georgios; Di Cai; Kappas, Arvid (2010): Sentiment Strength Detection in Short Informal Text. In *Journal of the American Society for Information Science and Technology* 61 (12), pp. 2544–2558.

Thomason, R. H. (1972): A Semantic Theory of Sortal Incorrectness. In *Journal of Philosophical Logic* 1 (2), pp. 209–258.

Thurman, Neil (2016): The Globalization of Journalism Online. A Transatlantic Study of News Websites and Their International Readers. In *Journalism: Theory, Practice & Criticism* 8 (3), pp. 285–307.

Tibshirani, Robert; Hastie, Trevor; Narasimhan, Balasubramanian; Chu, Gilbert (2003): Class Prediction by Nearest Shrunken Centroids, with Applications to DNA Microarrays. In *Statistical Science* 18 (1), pp. 104–117.

Tjong Kim Sang, Erik F. (2002): Introduction to the CoNLL-2002 Shared Task: Language-Independent Named Entity Recognition. In: Proceedings of the 6th Conference on Natural language Learning (CoNLL 2002). Taipei, Taiwan, August 26-30. Morristown, NJ, USA: Association for Computational Linguistics, pp. 1–4.

Tjong Kim Sang, Erik F.; Meulder, Fien de (2003): Introduction to the CoNLL-2003 Shared Task. Language-Independent Named Entity Recognition. In: Proceedings of the Seventh Conference on Natural Language Learning (CoNLL-2003). Edmonton, Canada, May 31-June 1: Edmonton, Canada, pp. 142–147.

Tolstoj, Aleksej N. (1956): O literature. Moskva: Sov. Pisatel'.

Tottie, Gunnel (1991): Negation in English Speech and Writing. A Study of Variation. San Diego: Academic Press.

Toutanova, Kristina; Klein, Dan; Manning, Christopher D.; Singer, Yoram (2003): Feature-rich Part-of-speech Tagging with a Cyclic Dependency Network. In: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology (NAACL 2003). Edmonton, Canada. Morristown, NJ, USA: Association for Computational Linguistics, pp. 173–180.

Toutanova, Kristina; Manning, Christopher D. (2000): Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-speech Tagger. In: Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics. Hong Kong, China, October 7-8. Morristown, NJ, USA: Association for Computational Linguistics, pp. 63–70.

Tsatsaronis, George; Panagiotopoulou, Vicky (2009): A Generalized Vector Space Model for Text Retrieval Based on Semantic Relatedness. In: Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop (EACL 2009). Athens, Greece, April 2. Athens, Greece: Association for Computational Linguistics, pp. 70–78.

Tsytsarau, Mikalai; Palpanas, Themis (2011): Towards a Framework for Detecting and Managing Opinion Contradictions. In: Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW 2011). Vancouver, BC, Canada, December 11: IEEE, pp. 1219–1222.

Tsytsarau, Mikalai; Palpanas, Themis; Denecke, Kerstin (2010): Scalable Discovery of Contradictions on the Web. In: Proceedings of the 19th international conference on World wide web (WWW 2010). Raleigh, NC, USA, April 26-30. New York, New York, USA: ACM Press, pp. 1195–1196.

Tsytsarau, Mikalai; Palpanas, Themis; Denecke, Kerstin (2011): Scalable Detection of Sentiment-Based Contradictions. In: Proceedings of the 1st International Workshop on Knowledge Diversity on the Web (DiversiWeb 2011), vol. 762. Hyderabad, India, March 28-April 1, pp. 9–16.

Tuchman, Gaye (1980): Making News. A Study in the Construction of Reality. First Free Press paperback ed. New York: Free Press.

Tulving, Endel (1972): Episodic and Semantic Memory. In Endel Tulving, Wayne Donaldson (Eds.): Organization of Memory. S.l.: Academic Press, pp. 382–402.

Tunstall, Jeremy (1996): Newspaper Power. The New National Press in Britain. Oxford: Oxford University Press.

Turian, Joseph; Ratinov, Lev; Bengio, Yoshua (2010): Word Representations. A Simple and General Method for Semi-supervised Learning. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010). Uppsala, Sweden, July 11-16, pp. 384–394.

Ungerer, Friedrich (2004): Ads as News Stories, News Stories as Ads. The Interaction of Advertisements and Editorial Texts in Newspapers. In *Text - Interdisciplinary Journal for the Study of Discourse* 24 (3).

Uryupina, Olga (2010): Corry: A System for Coreference Resolution. In: Proceedings of the 5th International Workshop on Semantic Evaluation (SemEval 2010). Los Angeles, CA, USA, July 15-16, pp. 100–103.

Uryupina, Olga; Zanoli, Roberto (2016): Preprocessing Technology. In Massimo Poesio, Roland Stuckardt, Yannick Versley (Eds.): Anaphora Resolution. Algorithms, Resources, and Applications. New York NY: Springer Berlin Heidelberg, pp. 209–236.

Uzquiano, Gabriel (2017): Quantifiers and Quantification. In Edward N. Zalta (Ed.): The Stanford Encyclopedia of Philosophy. Spring 2017: Metaphysics Research Lab, Stanford University.

van den Broek, Paul (1994): Comprehension and Memory of Narrative Texts. Inferences and Coherence. In Morton Ann Gernsbacher (Ed.): Handbook of Psycholinguistics. San Diego: Academic Press, pp. 539–588.

van den Broek, Paul; Rapp, David N.; Kendeou, Panayiota (2011): Integrating Memory-Based and Constructionist Processes in Accounts of Reading Comprehension. In *Discourse Processes* 39 (2-3), pp. 299–316.

van der Sandt, Rob A. (1992): Presupposition Projection as Anaphora Resolution. In *Journal of Semantics* 9 (4), pp. 333–377.

van der Wouden, Ton (1997): Negative Contexts. Collocation, Polarity and Multiple Negation. London, New York: Routledge.

van der Wouden, Ton; Zwarts, Frans (1993): A Semantic Analysis of Negative Concord. In *SALT* 3, p. 202.

van der Wurff, R. (2005): Impacts of the Internet on Newspapers in Europe. Conclusions. In *International Communication Gazette* 67 (1), pp. 107–120.

van Dijk, Teun A. (1988): News as Discourse. Hillsdale, NJ, Hove: L. Erlbaum Associates.

van Dijk, Teun A.; Kintsch, Walter (1983): Strategies of Discourse Comprehension. New York, NY: Academic Press.

van Valin, Robert D.; LaPolla, Randy J. (1997): Syntax. Structure, meaning, and function. Cambridge: Cambridge University Press.

Vapnik, Vladimir (1995): The Nature of Statistical Learning Theory. New York, London: Springer.

Varma, Vasudeva; Bharat, Vijay; Kovelamudi, Sudheer; Bysani, Praveen; GSK, Santosh; Kumar N., Kiran et al. (2009): IIIT Hyderabad at TAC 2009. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Varma, Vasudeva; Pingali, Prasad; K, Rahul; Krishna, Sai; Ganesh, Surya; Sarvabhotla, Kiran et al. (2008): IIIT Hyderabad at TAC 2008. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Vater, Heinz (2005): Referenz-Linguistik. München: Fink.

Velde, Roger G. van de (1989): Man, Verbal Text, Inferencing, and Coherence. In Wolfgang Heydrich, Fritz Neubauer, Jańos S. Petöfi, Emil Sözer (Eds.): Connexity and Coherence. Analysis of Text and Discourse. Reprint 2012. Berlin/Boston: De Gruyter, pp. 174–217.

Vendler, Zeno (1967): Facts and Events. In Zeno Vendler (Ed.): Linguistics in Philosophy. Ithaca: Cornell University Press, pp. 12–146.

Vendryés, Joseph (1950): Sur la négation abusive. In *Bulletin de la société de linguistique de Paris* 46, pp. 1–18.

Vennemann, Theo (1975): Topic, Sentence Accent, and Ellipsis: A Proposal or Their Formal Treatment. In Edward L. Keenan (Ed.): Formal Semantics of Natural Language. Cambridge: Cambridge University Press, pp. 313–328.

Verhagen, Marc; Pustejovsky, James (2008): Temporal Processing with the TARSQI Toolkit. In: Proceedings of the 22nd International Conference on Computational Linguistics, Posters (COLING 2008). Manchester, UK, August 18-22.

Versley, Yannick; Björkelund, Anders (2016): Off-the-Shelf Tools. In Massimo Poesio, Roland Stuckardt, Yannick Versley (Eds.): Anaphora Resolution. Algorithms, Resources, and Applications. New York NY: Springer Berlin Heidelberg, pp. 237–266.

Versley, Yannick; Poesio, Massimo; Ponzetto, Simone Paolo (2016): Using Lexical and Encyclopedic Knowledge. In Massimo Poesio, Roland Stuckardt, Yannick Versley (Eds.): Anaphora Resolution. Algorithms, Resources, and Applications. New York NY: Springer Berlin Heidelberg, pp. 393–429.

Versley, Yannick; Ponzetto, Simone Paolo; Poesio, Massimo; Eidelman, Vladimir; Jern, Alan; Smith, Jason et al. (2008): BART. In: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies Demo Session (HLT 2008). Columbus, OH, USA, June 16. Morristown, NJ, USA: Association for Computational Linguistics, pp. 9–12.

Vieira, Renata; Poesio, Massimo (2000): An Empirically Based System for Processing Definite Descriptions. In *Computational Linguistics* 26 (4), pp. 539–593.

Vincze, Veronika; Szarvas, György; Farkas, Richárd; Móra, György; Csirik, János (2008): The BioScope Corpus. Biomedical Texts Annotated for Uncertainty, Negation and Their Scopes. In *BMC bioinformatics* 9 Suppl 11, S9.

Voorhees, Ellen M. (2008): Contradictions and Justifications. Extensions to the Textual Entailment Task. In: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL 2008). Columbus, OH, USA, June 15-20, pp. 63–71.

Vosniadou, Stella; Pearson, P. David; Rogers, Theresa (1988): What Causes Children's Failures to Detect Inconsistencies in Text? Representation versus Comparison Difficulties. In *Journal of Educational Psychology* 80 (1), pp. 27–39.

Wang, Rui; Neumann, Günter (2007): Recognizing Textual Entailment Using Sentence Similarity Based on Dependency Tree Skeletons. In: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing (RTE 2007). Prague, Czech Republic, June 28-29, pp. 36–41.

Wang, Rui; Neumann, Günter (2008): An Accuracy-Oriented Divide-and-Conquer Strategy for Recognizing Textual Entailment. In: Proceedings of the Fourth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2008). Gaithersburg, MD, USA, November 17-19.

Wang, Rui; Zhang, Yi; Neumann, Günter (2009): A Joint Syntactic-Semantic Representation for Recognizing Textual Relatedness. In: Proceedings of the Fifth PASCAL Challenges Workshop on Recognising Textual Entailment (RTE 2009). Gaithersburg, MD, USA, November 16-17.

Wartena, Christian; Anjewierden, Anjo; van Dieten, William (2006): Using Patterns over Ontology Terms to Detect Inconsistencies in Large Text Collections. In: Proceedings of the International Workshop on Ontologies in Text Technology: Approaches to Extract Semantic Knowledge from Syntactic Information (OTT 2006). Osnabrück, Germany, August, pp. 107–112.

Wassenburg, Stephanie I.; Beker, Katinka; van den Broek, Paul; van der Schoot, Menno (2015): Children's Comprehension Monitoring of Multiple Situational Dimensions of a Narrative. In *Reading and writing* 28 (8), pp. 1203–1232.

Wedin, Michael (2004): Aristotle on the Firmness of the Principle of Non-Contradiction. In *Phronesis* 49 (3), pp. 225–265.

Weimer, Wolfgang (2005): Logisches Argumentieren. Stuttgart: Reclam.

Weischenberg, Siegfried (2001): Nachrichten-Journalismus. Anleitungen und Qualitäts-Standards für die Medienpraxis. 1. Aufl. Wiesbaden: Westdeutscher Verlag.

White, Peter Robert Rupert (1997): Death, Disruption and the Moral Order: The Narrative Impulse in Mass-Media 'Hard News' Reporting. In Frances Christie, J. R. Martin (Eds.): Genre and Institutions. Social Processes in the Workplace and School. London: Cassell, pp. 101–133.

Wiebe, Janyce M. (1994): Tracking Point of View in Narrative. In *Computational Linguistics* 20 (2), pp. 233–287.

Wiebe, Janyce M.; Bruce, Rebecca F.; O'Hara, Thomas P. (1999): Development and Use of a Gold-Standard Data Set for Subjectivity Classifications. In: Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics (ACL 1999). College Park, MD, USA, June 20-26. Morristown, NJ, USA: Association for Computational Linguistics, pp. 246–253.

Wiebe, Janyce M.; O'Hara, Thomas P.; Öhrström-Sandgren, Thorsten; McKeever, Kenneth J. (1998): An Empirical Approach to Temporal Reference Resolution. In *Journal of Artificial Intelligence* 9 (1), pp. 247–294.

Wiebe, Janyce M.; Wilson, Theresa; Bruce, Rebecca; Bell, Matthew; Martin, Melanie (2004): Learning Subjective Language. In *Computational Linguistics* 30 (3), pp. 277–308.

Wiebe, Janyce M.; Wilson, Theresa; Cardie, Claire (2005): Annotating Expressions of Opinions and Emotions in Language. In *Language Resources and Evaluation* 39 (2-3), pp. 165–210.

Williams, John N. (1981): Inconsistency and Contradiction. In *Mind* 90, pp. 600–602.

Wilson, Theresa; Wiebe, Janyce; Hwa, Rebecca (2006): Recognizing Strong and Weak Opinion Clauses. In *Computational Intelligence* 22 (2), pp. 73–99.

Winograd, Peter; Johnston, Peter (1982): Comprehension Monitoring and the Error Detection Paradigm. In *Journal of Reading Behavior* 14 (1), pp. 61–76.

Wittgenstein, Ludwig (1953): Philosophical Investigations. New York: Macmillan.

Wolfram, Sybil (1989): Philosophical Logic. An Introduction. London: Routledge.

Xu, Jinxi; Croft, W. Bruce (1998): Corpus-based Stemming Using Cooccurrence of Word Variants. In *ACM Transactions on Information Systems* 16 (1), pp. 61–81.

Yalcin, Seth (2007): Epistemic Modals. In *Mind* 116 (464), pp. 983–1026.

Yates, Alexander; Etzioni, Oren (2007): Unsupervised Resolution of Objects and Relations on the Web. In: Proceedings of the Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (NAACL-HLT 2007). Rochester, NY, USA, April 22-27.

Young, Peter; Lai, Alice; Hodosh, Micah; Hockenmaier, Julia (2014): From Image Descriptions to Visual Denotations: New Similarity Metrics for Semantic Inference over Event Descriptions. In *Transactions of the Association for Computational Linguistics* 2, pp. 67–78.

Yule, George (1996): The study of Language. 2nd ed. Cambridge: Cambridge University Press.

Yule, George (2017): The study of Language. Sixth edition. Cambridge, UK: Cambridge University Press.

Zanoli, Roberto; Colombo, Silvia (2016): A Transformation-driven Approach for Recognizing Textual Entailment. In *Natural Language Engineering* 23 (04), pp. 507–534.

Zhu, Qiaoming; Li, Junhui; Wang, Hongling; Zhou, Guodong (2010): A Unified Framework for Scope Learning via Simplified Shallow Semantic Parsing. In: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (EMNLP 2010). Cambridge, MA, USA, October 9-11, pp. 714–724.

Zimmer, Karl E. (1964): Affixal Negation in English and Other Languages. An Investigation of Restricted Productivity. London: Clowes.

Appendix A. Survey 1: An Example of a Questionnaire

Studie zu textuellen Widersprüchen in Nachrichtentexten

Name: Englischkenntnisse:
(Muttersprachlerin / fließend / sehr gut / gut / Grundkenntnisse)

Vielen Dank, dass Sie an der vorliegenden Studie teilnehmen. Im Rahmen meiner Promotion arbeite ich an der Entwicklung einer Methode für die automatische Erkennung von widersprüchlichen Aussagen in englischen Texten. Sie können mich dabei unterstützen, indem Sie die Texte in diesem Fragebogen sorgfältig lesen und die in diesen Texten möglich enthaltenen Widersprüche erkennen.

Bitte lesen Sie die folgenden Hinweise, bevor Sie mit der Bearbeitung des Fragebogens beginnen.

- Im Folgenden finden Sie Nachrichtentexte zu zwei bekannten Weltereignissen. Zu jedem Weltereignis gibt es jeweils 4 bis 8 Texte, die aus unterschiedlichen Quellen stammen. Die Sätze in Texten je Weltereignis sind fortlaufend nummeriert.

- Lesen Sie die Texte bitte sorgfältig durch. Bei auffallenden widersprüchlichen Aussagen **im Text** oder **zwischen den Texten** tragen Sie die Nummern der Sätze, die diese Aussagen enthalten, in die dafür vorgesehene Tabelle (s. Vorlage zum Thema) unter „Sätze mit widersprüchlichen Aussagen“ ein. Bitte begründen Sie kurz Ihre Entscheidung.

- Bitte beachten Sie, dass manche Widersprüche komplex sind und sich über mehrere Sätze erstrecken können. In diesem Fall tragen Sie alle dazu gehörigen Sätze in die Tabelle ein.

Beispiel:

Sie haben drei Aussagen gefunden, die zueinander im Widerspruch stehen. Die erste Aussage wird im Satz 23 geäußert, die zweite im Satz 198. Die dritte Aussage ihrerseits ist komplex und befindet sich in den Sätzen 47, 48, 49. Tragen Sie in die Tabelle daher *23 / 198 / 47, 48, 49* ein. Vergessen Sie bitte nicht, Ihre Entscheidung kurz zu begründen.

- Hilfsmittel aller Arten (Fremd- und Bedeutungswörterbücher, Wikipedia etc.) dürfen benutzt werden.

- Einige Widersprüche werden auf der nachfolgenden Seite dargestellt.

- *Bei Fragen zum Aufbau und Inhalt der Studie sowie bei auffallenden Fehlern kontaktieren Sie mich bitte per Email unter natali.karlova-bourbonus@zmi.uni-giessen.de.*

Den ausgefüllten Fragebogen können Sie bei mir persönlich abgeben oder per Post bis **zum 30.11.2014** zusenden an

Professur für Angewandte Sprachwissenschaft und Computerlinguistik
z.Hd. Natali Karlova-Bourbonus
Otto-Behaghel-Straße 10D, Büro D406
35394 Gießen

Bitte beachten Sie, erst nach vollständiger Bearbeitung und einer damit verbundenen Abgabe des Fragebogens erhalten Sie einen Gutschein im Wert von 5 Euro!

Beispiele der widersprüchlichen Aussagen (Auswahl)

- **Enthalten widersprüchliche Fakten**

- Russia´s President Putin has spoken to Western leaders to emphasize "the extreme importance of not allowing a further escalation of violence", the Kremlin said.

- Russia has denied any involvement with the airports takeover but confirmed its armoured vehicles had been on the move around Crimea.

- **Entstehen durch Verwendung von Negationen**

- […]. Chinese citizens – 154, […], Russia, Italy, the Netherlands and Austria – one each.

- The foreign ministry in Rome said no Italian was on the plane either, despite the inclusion of Maraldi´s name on the list.

- **Entstehen durch Gebrauch von Antonymen**

- Capital punishment is a catalyst for more crime.

- Capital punishment is a deterrent to crime.

- **Enthalten widersprüchliche numerische Angaben**

- The Boeing B777-200 aircraft was carrying 227 passengers, including two children, and 12 crew members.

- The plane was carrying five children under 5 years old, the airline said.

- **Widersprüchlichkeit aufgrund der grammatischen Funktionen der Wörter**

- Jacques Santer *succeeded* Jacques Delors as president of the European Commission in 1995.

- Delors *succeeded* Santer in the presidency of the European Commission.

- **Hintergrundwissen nötig, um diese Art des Widerspruchs zu erkennen**

- The jet carried 227 passengers from 14 countries, mainly China and Malaysia, and a crew of 12, all Malaysian nationals, the carrier said in a statement.

- The passengers were of 14 different nationalities, Mr. Yahya said. **(country vs. nationality)**

Vorlage zum Thema “Malaysia Airlines Flight MH370 Missing”

Sätze mit widersprüchlichen Aussagen	Kurze Begründung

Titel:	Malaysia Airlines flight 'presumed crashed'
Thema:	Malaysia Airlines Flight MH370 Missing
Quelle / Datum:	Al Jazeera, 09.03.2014

(1) Air and marine search for plane missing on way to Beijing with 239 people onboard continues in sea south of Vietnam.

(2) A Malaysia Airlines flight carrying 227 passengers and 12 crew has gone missing over the South China Sea, presumed crashed.

(3) The airline on Saturday said search and rescue teams from countries closest to the missing plane's flight path had been sent to scour a large area near its last known location.

(4) Vietnam said its rescue planes spotted two large oil slicks in the sea and it was sending boats to the area.

(5) "Two of our aircraft sighted two oil slicks around 15 to 20km long, running parallel, around 500 metres apart from each other," the army's deputy chief of staff, Vo Van Tuan, told state-run VTV.

(6) A crash, if confirmed, would mark the United States-built Boeing 777-200ER airliner's deadliest incident since entering service 19 years ago.

(7) Malaysia's flag carrier said flight MH370 disappeared, without giving a distress signal, at 2:40am local time on Saturday (18:40 GMT Friday), about two hours after leaving Kuala Lumpur International Airport.

(8) It had been due to arrive in Beijing at 6:30am local time on Saturday (22:30 GMT Friday).

(9) Passports stolen

(10) The flight was carrying 154 people from China or Taiwan, 38 Malaysians, seven Indonesians, six Australians and five Indians, the airline said.

(11) There were also three US citizens, four from France, two passengers each from New Zealand, Ukraine, and Canada, and one each from Italy, the Netherlands, Russia and Austria, the airline said in a statement.

(12) However, Foreign Ministry officials in Rome and Vienna later said names of two nationals listed as passengers matched passports reported stolen in Thailand.

(13) The Austrian, whose passport was stolen two years ago, was found safe at home, a ministry spokesman said.

(14) Italian news agency ANSA said Luigi Maraldi called home after hearing reports that an Italian with his name was aboard the plane.

(15) Al Jazeera's Rob McBride, reporting from Beijing, said that it is a very hard situation for the airline as it does not have the visual confirmation that its plane has crashed.

(16) "The company does not want to say so until it has confirmation," he said.

(17) Al Jazeera’s Florence Looi, reporting from Kuala Lumpur, said that the search teams concentrate rescue efforts on the area where contact was last made with the aircraft.

(18) "There are more than a dozen Malaysian planes involved in the search and rescue mission and about nine ships from the same country.

(19) Singapore and Vietnam are also involved in the mission.

(20) And the US is sending two ships," she said.

(21) The airline said it was working with authorities in the search efforts to locate the aircraft.

(22) No distress signal

(23) Ross Aimer, a former pilot with United Airlines, told Al Jazeera it was highly unusual that air traffic control would lose contact with an aircraft without communication from the crew.

(24) "The fact that there was absolutely no distress signal is very disturbing.

(25) This is almost unprecedented that we lose an aircraft in such a way …

(26) In that area of the world, over Vietnam, there is sporadic radar coverage to begin with," he said.

(27) A report by China's Xinhua news agency said contact was lost with the plane while it was near Vietnamese airspace.

(28) The airline's Kuala Lumpur-Beijing route passes roughly over the Indochinese Peninsula.

Titel:	Missing Malaysia plane: 'Oil slick seen'
Thema:	Malaysia Airlines Flight MH370 Missing
Quelle / Datum:	BBC, 08.03.2014

(29) A multinational team is searching the sea off south Vietnam, in the hope of finding a Malaysia Airlines flight that has been missing for 24 hours.

(30) A Vietnamese search plane saw two possible oil slicks in the area, although there was no confirmation they were related to the disappearance.

(31) Flight MH370 had 239 people on board, en route to Beijing from Kuala Lumpur.

(32) Two-thirds of the passengers were from China, while others were from elsewhere in Asia, North America and Europe.

(33) It has been reported that two passengers who were listed on the plane's manifest - an Italian and an Austrian - were not actually on the flight.

(34) They both reportedly had their passports stolen in Thailand.

(35) Asked whether terrorism was suspected as a reason for the plane's disappearance, Malaysian Prime Minister Najib Razak said: "We are looking at all possibilities, but it is too early to make any conclusive remarks."

(36) A senior US official told NBC News: "We are aware of the reporting on the two stolen passports.

(37) We have not determined a nexus to terrorism yet, although it's still very early, and that's by no means definitive."

(38) US help

(39) Flight MH370 vanished at 17:30 GMT Friday (01:30 local time Saturday).

(40) The plane reportedly went off the radar south of Vietnam, and according to Malaysian Airlines, it last had contact with air traffic controllers 120 nautical miles off the east coast of the Malaysian town of Kota Bharu.

(41) Distraught relatives and loved ones of those on board are being given assistance at both the arrival and departure airports.

(42) Malaysia Airlines chief executive Ahmad Jauhari Yahya said the focus was on helping the families of those missing.

(43) He said that 80% of the families had been contacted.

(44) The passengers were of 14 different nationalities, Mr Yahya said.

(45) Among them were 153 Chinese nationals, 38 Malaysians, seven people from Indonesia and six from Australia.

(46) Malaysia and Vietnam have both sent planes and naval vessels to search for the missing flight, and the US is sending the USS Pinckney, an Arleigh Burke-class guided-missile destroyer, which could be in the area within 24 hours.

(47) Territorial disputes over the South China Sea were set aside temporarily as China dispatched two maritime rescue ships and the Philippines deployed three air force planes and three navy patrol ships.

(48) Singapore is also involved, while Vietnam sent aircraft and ships and asked fishermen in the area to report any suspected sign of the missing plane.

(49) The pilot was Capt Zaharie Ahmad Shah, 53, who joined Malaysia Airlines in 1981, Mr Yahya said.

(50) Friends and relatives expecting to meet passengers from the flight in Beijing were instructed to go to a nearby hotel where officials were meant to be on hand to provide support.

(51) "They should have told us something before now," a visibly distressed man in his thirties told AFP news agency at the hotel.

(52) "They are useless," another young man said of the airline.

(53) "I don't know why they haven't released any information."

(54) In Kuala Lumpur, Hamid Ramlan, a 56-year-old police officer, said his daughter and son-in-law had been on the flight for an intended holiday in Beijing.

(55) "My wife is crying," he said.

(56) "Everyone is sad.

(57) My house has become a place of mourning.

(58) This is Allah's will.

(59) We have to accept it."

(60) Malaysia's national carrier is one of Asia's largest, flying nearly 37,000 passengers daily to some 80 destinations worldwide.

(61) The route between Kuala Lumpur and Beijing has become more and more popular as Malaysia and China increase trade, says the BBC's Jennifer Pak in Kuala Lumpur.

(62) The Boeing 777 had not had a fatal crash in its 20-year history until an Asiana plane came down at San Francisco airport in July of last year.

(63) Three teenage girls from China died in that incident.

(64) Aviation expert David Learmount told the BBC that passenger planes today "are incredibly reliable and you do not get some sudden structural failure in flight - it just doesn't happen".

Titel:	No signs of crash for Malaysian flight found: transport official
Thema:	Malaysia Airlines Flight MH370 Missing
Quelle / Datum:	China News, 08.03.2014

(65) Malaysian transport authorities said that no signs had been found that a Malaysia Airlines flight, carrying 239 passengers and crew had crashed, Transport Minister Hishamuddin Hussein said in Kuala Lumpur Saturday.

(66) He said no sign of any plane wreckage was found and denied earlier media reports that the plane had crashed south of an island off Vietnam.

(67) "We are doing everything in our power to locate the plane, and doing everything we can to ensure every possible angle has been addressed," Hishamuddin told reporters near the Kuala Lumpur International Airport.

(68) "We are looking for accurate information from the Malaysian military.

(69) They are waiting for information from the Vietnamese side."

(70) Meanwhile, the Malaysia Airlines said in a latest statement that the company is working with international authorities on the search and rescue mission of its flight MH370 that went missing on Saturday morning with 239 people on board.

(71) "Our team is currently calling family members of passengers to keep them updated on the situation and our focus now is to work with the emergency responders and the authorities.

(72) We are sending a MH team to support the families of passengers at Beijing.

(73) The airline will continue to publish regular updates on the situation.

(74) Flight MH 370, operating a Boeing B777-200 aircraft, departed Kuala Lumpur at 0:21 am local time (1621GMT) and was expected to land in the Chinese capital at 6:30 am (2230GMT) the same day.

(75) The flight was piloted by Captain Zahaire Ahmad Shah, a Malaysian aged 53.

(76) He has a total flying hours of 18,365 hours.

(77) He joined Malaysia Airlines in 1981.

(78) Fariq Ab. Hamid, 27, also a Malaysian, served as the first officer of the flight.

(79) With a total of flying hours of 2.763, he joined Malaysian Airlines in 1981, the airlines said.

(80) Earlier, Vietnam's Tuoi Tre (Youth) newspaper quoted Rear Admiral Ngo Van Phat, political commissar of the Fifth Naval Region as saying that the missing aircraft had crashed into waters off Vietnam's southern Phu Quoc Island.

(81) The allegation was denied by the Malaysian authorities.

Titel:	Malaysia Airlines Loses Contact With Jet Carrying Over 200
Thema:	Malaysia Airlines Flight MH370 Missing
Quelle / Datum:	New York Times, 07.03.2014

(82) Malaysia Airlines announced Saturday morning that it had lost contact five hours earlier with one of its flights, which was carrying at least 239 people to Beijing from Kuala Lumpur, and had activated a search-and-rescue team.

(83) The plane, a Boeing 777-200 operating as Flight MH370, took off at 12:41 a.m.

(84) Air traffic control in Subang, a suburb of Kuala Lumpur, lost contact with the plane almost two hours later, at 2:40 a.m.

(85) The plane was scheduled to land at 6:30 a.m. in Beijing, but there was no further word on its fate by early Saturday afternoon.

(86) Vietnamese officials told local news media that the aircraft had never reached the air traffic control region for Ho Chi Minh City, formerly Saigon, after the plane was supposed to have passed over the ocean between northern Malaysia and southernmost Vietnam.

(87) Malaysia Airlines said the flight had 227 passengers aboard, including two infants, and a crew of 12.

(88) Airline staff members have begun contacting the families of passengers and crew members.

(89) “Our thoughts and prayers are with all affected passengers and crew and their family members,” the airline’s statement said.

(90) The arrival board at Beijing Airport listed the Malaysia Airlines flight that lost contact with air traffic controllers on Saturday.

(91) Chinese officials expressed immediate concern.

(92) “We are extremely worried upon hearing this news,” Qin Gang, the spokesman for the Chinese Foreign Ministry, said in a statement.

(93) “We are currently in contact with relevant parties and are doing what we can to understand and confirm relevant circumstances.”

(94) He added that the Foreign Ministry, the Chinese Embassy in Malaysia and the Chinese Embassy in Vietnam had begun emergency procedures.

(95) If the flight was traveling in a straight line, it would have been traveling north up the entire length of Vietnam and Vietnam’s coastal waters.

(96) VnExpress, an online newspaper in Vietnam, quoted Dinh La Thang, Vietnam’s transportation minister, as saying that the flight had not reached the so-called flight information region for Ho Chi Minh City.

(97) Chinese air traffic control authorities said the plane had not entered airspace that they control or established communications with Chinese air traffic control, according to the state-owned China Central Television.

(98) Malaysia Airlines said more than 150 Chinese and four Americans were aboard the plane.

(99) In the terminal at Beijing International Airport where Flight MH370 had been scheduled to arrive, a woman burst into tears while on a telephone.

(100) Liu Meng, 26, said he had been at the airport since shortly before the flight's scheduled arrival time, waiting for his boss to return from a business trip.

(101) The boss's relatives had been calling Mr. Liu with questions, and, he said, he had nothing to tell them.

(102) A Malaysian man who gave only his surname, Zhang, said he had been waiting at the airport for two Malaysian friends on the flight, but the airport authorities had told him only that the flight had been delayed; he learned news of the aircraft's disappearance from reading about it online.

(103) There have been two previous crashes of Boeing 777s.

(104) Last July 6, an Asiana plane came in too slow and at too low an altitude and crash-landed at San Francisco International Airport.

(105) Three people were killed and several others suffered serious permanent injuries.

(106) So far it does not appear that there was a mechanical problem with that aircraft.

(107) In January 2008, a British Airways 777 came in short of the runway at Heathrow in London.

(108) Both engines failed.

(109) Problem was traced to icing in the fuel system.

(110) Nobody was killed.

Titel:	For Families of Missing on Airliner, Memories Mix With Fading Hope
Thema:	Malaysia Airlines Flight MH370 Missing
Quelle / Datum:	New York Times, 08.03.2014

(111) A Canadian couple returning from vacation in Vietnam.

(112) An American who worked in Asia for IBM.

(113) A group of Chinese calligraphers who had attended an exhibition in Malaysia.

(114) All of them were aboard Malaysia Airlines Flight MH370, which remained unaccounted for on Saturday, many hours after it should have landed at dawn in Beijing with 239 people on board, most of them from China.

(115) By Saturday night, the families of the passengers had few answers about what happened and dwindling hope that they would see their loved ones again.

(116) One passenger was Philip Wood, 50, an IBM employee who was living in the Malaysian capital, Kuala Lumpur, where the flight originated.

(117) “We’re all sticking together,” his father, Aubrey Wood, said from his home in Keller, Tex.

(118) “What can you do? What can you say?”

(119) Philip Wood, who previously lived in Beijing, has two sons in Texas.

(120) He had followed in his father’s footsteps when he joined IBM, from which his father retired at the end of his career.

(121) The State Department confirmed that there were three Americans on board.

(122) The two other Americans listed on the flight manifest were Nicole Meng, 4, and Yan Zhang, 2.

(123) It was unclear whether the children were traveling with parents from China or another country, living in the United States when they were born, or traveling with American parents with dual citizenship.

(124) The two Canadian citizens on the plane were Muktesh Mukherjee, 42, and Xiaomo Bai, 37, a married couple who had left two young sons at home in Beijing while they vacationed in Vietnam.

(125) Mr. Mukherjee worked in Beijing for Xcoal Energy & Resources.

(126) The company’s chief executive, Ernie Thrasher, called him “a dear friend, colleague and member of the Xcoal family.”

(127) The couple had “two wonderful little boys,” said Matthew McConkey, a close friend of Mr. Mukherjee, whom he had seen recently in Beijing.

(128) “A big group of us went out,” said Mr. McConkey.

(129) “He was just always happy.

(130) Life had been good to him.”

(131) Mr. McConkey said he was relieved that their children were safe in Beijing, but he added, “This is just a nightmare.”

(132) The couple posted photos on social media last week from a resort on the coast of Vietnam.

(133) They often shared photos of their children on Facebook — including one a month ago of the boys making snow angels outside an apartment building called Central Park in Beijing.

(134) A group of as many as 24 painters and calligraphers were returning from an exhibition and a cultural exchange conference in Kuala Lumpur.

(135) The conference was dedicated to the “Chinese Dream” and intended to celebrate the 40th anniversary of diplomatic relations between China and Malaysia.

(136) The Sichuan provincial government said Zhang Jinquan, 72, a well-known calligrapher, was on the plane, and the manifest listed Meng Gaosheng, 64, vice chairman of the China Calligraphic Artists Association.

(137) One of the younger members of the delegation was Maimaitijiang Abula, 35, an art teacher at a college in Kashgar.

(138) A friend, Kuerbanjiang Saimaiti, described him as a talented oil painter who once confided that he wanted to spend “a lifetime on painting well” and recently completed advanced studies at an art academy in Beijing.

(139) At the Kuala Lumpur airport, a grief-stricken relative of Chng Mei Ling screamed uncontrollably as airline employees escorted him out of the terminal.

(140) “Be truthful about this!” said Koon Chim Wa, the relative, whose booming voice echoed through the cavernous terminal.

(141) “They say they don’t know where the plane is,” Mr. Koon said, his hands and body shaking.

(142) “Is this a joke?”

(143) His niece, Ms. Chng, a Malaysian engineer working at a company in Pennsylvania, was on her way to the United States, via Beijing, Mr. Koon said.

(144) In Beijing, Lu Jiang, 32, told The China Daily that her neighbor was on the list of passengers.

(145) “I saw her name and the name of her husband and her 1-year-old baby on the missing passenger’s list,” Ms. Lu said.

(146) “I never thought this would happen.

(147) God bless them.”

Titel:	Malaysian plane crashed off Vietnam coast: state media
Thema:	Malaysia Airlines Flight MH370 Missing
Quelle / Datum:	Reuters, 08.03.2014

(148) A Malaysia Airlines flight carrying 227 passengers and 12 crew crashed into the sea 153 miles off the coast of Vietnam's Tho Chu island on Saturday, according to a Vietnamese navy officer quoted by state media.

(149) "At the moment there are no Vietnamese navy boats in that area so we have to ask boats from Phu Quoc island to be prepared for rescue," Admiral Ngo Van Phat told the website of Tuoi Tre news.

(150) Tho Chu and Phu Quoc lie to the southwest of southern Vietnam.

(151) The admiral could not be reached by telephone.

(152) It was not immediately clear how he knew where the plane had crashed or whether wreckage had been spotted.

(153) The plane last had contact with air traffic controllers 120 nautical miles off the east coast of the Malaysian town of Kota Bharu, the airline said.

Titel:	Missing plane may have turned back before disappearing: official
Thema:	Malaysia Airlines Flight MH370 Missing
Quelle / Datum:	Xinhua, 09.03.2014

(154) A China-bound Malaysia Airlines plane may have turned back before it went missing from radar screen, a Malaysian civil aviation official said Sunday.

(155) The Malaysian authorities are also investigating two passengers that had used false passports to board the plane, department of civil aviation director general Azharuddin Abdul Rahman said.

(156) A Malaysia Airlines Boeing 777-200 passenger plane with 239 people onboard, including more than 150 Chinese, lost contact with air traffic controllers en route from Kuala Lumpur to Beijing early Saturday.

(157) There are still no confirmed information about the fate of the plane after about 36 hours passed.

Vorlage zum Thema “Typhoon Haiyan Hits Philippine”

Sätze mit widersprüchlichen Aussagen	Kurze Begründung

Titel:	Scores dead after typhoon hits Philippines
Thema:	Typhoon Haiyan hits Philippine
Quelle / Datum:	Al Jazeera, 09.11.2013

(1) Officials warn of "innumerable casualties" as one of the world's strongest storms on record hits southeast Asia.

(2) One of the strongest storms on record has slammed into the central Philippines, killing at least 100 people, forcing hundreds of thousands from their homes and knocking out power and communications in several provinces.

(3) Typhoon Haiyan left the Philippines early on Saturday on a path toward Southeast Asia, the US National Oceanic and Atmospheric Administration tweeted.

(4) Forecasters said the storm was expected to pick up renewed strength over the South China Sea on its way towards Vietnam.

(5) As the storm left on Saturday morning, thousands of Philippine soldiers raced to reach isolated communities that were devastated by the typhoon, as reports emerged of corpses lining roads and people being swept out to sea.

(6) More than 100 bodies were lying in the streets of one Philippine city that was hit by Haiyan, an aviation authority chief said on Saturday.

(7) Military sources in Leyte province, once of the worst hit, told Al Jazeera that an initial survey showed there were "innumerable casualties" there.

(8) "We have reports of collapsed buildings, houses flattened to the ground, storm surges and landslides," Philippine Red Cross chief Gwendolyn Pang told the AFP news agency.

(9) Nichola Jones, a representative of the International Federation of the Red Cross and Red Crescent Societies, told Al Jazeera that her organisation had received reports of "widespread devastation" and flooding in Leyte.

(10) Initial reports from the worst-affected areas indicated that communications and transportation networks had been crippled, and many homes washed away.

(11) Haiyan smashed into the eastern provinces of Leyte and Samar with maximum sustained winds of around 315 kilometres per hour.

(12) Death toll to rise

(13) The death toll was expected to rise, with authorities unable to immediately contact the worst-affected areas.

(14) "The winds were so strong that they flattened all the banana plants around the house," university student Jessa Aljibe, 19, told AFP by telephone from the Samar city of Borongan shortly after Haiyan made landfall.

(15) Power and communications in the three large island provinces of Samar, Leyte and Bohol were almost completely down but the government and telephone service providers promised to restore them within 24 hours, the Reuters news agency reported.

(16) Authorities warned that more than 12 million people were at risk, including residents of Cebu City, which has a population of about 2.5 million, and areas still reeling from a deadly 2011 storm and a 7.2-magnitude quake last month.

(17) Thousands of people were evacuated from villages in the country's central regions, including a province devastated by an earthquake last month.

(18) David Carden, who heads the UN Office for the Coordination of Humanitarian Affairs, spoke to Al Jazeera about the severity of the storm.

(19) "We have heard there had been high storm surges, in some areas as high as the second floor of houses," he said.

(20) "In the areas directly affected there is no power.

(21) People have no water and [there has been] significant damage to shelters.

(22) Information is still coming in and we are, of course, very concerned."

(23) War-like preparations

(24) President Benigno Aquino III gave warning to people to leave high-risk areas, including 100 coastal communities where forecasters said the storm surge could reach up to seven metres.

(25) He urged seafarers to stay in port.

(26) Aquino also assured the public of war-like preparations: three C-130 air force planes, 32 helicopters and 20 ships were on standby.

(27) "No typhoon can bring Filipinos to their knees if we're united," he said in a televised address.

(28) Edgardo Chatto, governor of Bohol island province in the central Philippines, where an earthquake in October killed more than 200 people, said soldiers, police and rescue units were helping displaced residents, including thousands staying in small tents, move to shelters.

(29) Bohol is not forecast to receive a direct hit but is expected to be battered by strong winds and rain, government forecaster Jori Loiz said.

Titel:	Monster typhoon Haiyan roars across Philippines
Thema:	Typhoon Haiyan hits Philippine
Quelle / Datum:	BBC, 08.11.2013

(30) One of the strongest typhoons ever to hit land has slammed the Philippines, forcing millions to take shelter.

(31) Packing sustained winds of up to 320 km/h (199mph), Typhoon Haiyan left at least four people dead, but it may be days before the full damage is known.

(32) The storm ripped apart buildings and triggered landslides as it ploughed across the country's central islands.

(33) Officials said more than 12 million people were at risk, but the storm's rapid passing could limit its impact.

(34) "We expect the level of destruction caused by Typhoon Haiyan to be extensive and devastating, and sadly we fear that many lives will be lost," said Save the Children's Philippines director Anna Lindenfors.

(35) The Philippines has experienced more than its fair share of super typhoons over the past decade, according to experts.

(36) There were at least three of these powerful events in nine of the 10 years between 2002 and 2012.

(37) The islands are unlucky, scattered along the world's most active typhoon belt where plentiful supplies of warm water and moist air provide the energy to kick start super storms.

(38) Despite these factors, Haiyan has shown a number of unusual features which have increased its strength.

(39) Normally the walls of the storm that rotate around the eye are replaced as it moves, often weakening the wind speed.

(40) In the case of Haiyan this hasn't happened.

(41) Another factor has been the speed of this typhoon.

(42) Going so quickly, it hasn't stirred up the waters ahead of it.

(43) Slower storms churn up the waters, causing an upwelling of colder water that usually takes the energy from the storm.

(44) However Haiyan has now lost energy over land and is expected to move on to Vietnam as a Category 3 Typhoon in the next few days.

(45) Eduardo del Rosario, head of the disaster response agency, told the Associated Press that early evacuations and the speed at which the typhoon swept across the Philippines, may have helped reduce its destructive potential.

(46) Lt Gen Roy Deveraturda, a military commander, echoed this view, telling the AP: "It has helped that the typhoon blew very fast in terms of preventing lots of casualties."

(47) Meteorologists had earlier warned that the storm could be as devastating as Typhoon Bopha in 2012, which ravaged parts of the southern Philippines and left at least 1,000 people dead.

(48) Haiyan - equivalent to a category five hurricane - is now heading towards Vietnam and southern China.

(49) The storm made landfall on the Philippines shortly before dawn, bringing gusts that reached 379 km/h (235 mph), waves as high as 15m (45ft) and up to 400mm (15.75 inches) of rain in places.

(50) There were reports of buildings being ripped apart, flash floods and landslides.

(51) Schools and offices were closed, while ferry services and local flights were suspended.

(52) Hospitals and soldiers were on stand-by for rescue and relief operations.

(53) Power and communication lines were also cut to some areas.

(54) Haiyan raged across Leyte and Samar, turning roads into rivers, and battered Cebu city, the country's second largest with a population of 2.5 million.

(55) The eye of the storm - known locally as Yolanda - passed well to the south of the capital Manila, but the city still felt its force.

(56) "The wind here is whistling.

(57) It's so strong and the heavy downpours are continuing," Mai Zamora, from the charity World Vision, in Cebu, told the BBC.

(58) "We've been hearing from my colleagues in [the city of] Tacloban that they've seen galvanised iron sheets flying just like kites."

(59) "It was frightening.

(60) The wind was so strong, it was so loud, like a screaming woman.

(61) I could see trees being toppled down," Liwayway Sabuco, a saleswoman from Catbalogan, a major city on Samar, told AFP news agency.

(62) Former BBC Manila correspondent Kate McGeown says that while reports are now coming in from some of the affected cities, there was still very little information from the countryside in large areas of the Visayas region such as Negros and Iloilo, and the island of Mindoro.

(63) There were reports of substantial damage even in areas that missed the worst of Haiyan, the 25th tropical storm to enter Philippine territory this year.

(64) "The storm was very strong - although Surigao City was not directly hit we experienced its fury early this morning," said Protestant pastor Diosdado Casera in Surigao City in northeast Mindanao .

(65) "The big buildings made of concrete were fine, but the houses made of wood and shingles and plywood have suffered a lot of damage, mainly to their roof."

(66) A spokesperson for the British Red Cross, Nichola Jones, who is in Tagbilaran on the central island of Bohol, said the typhoon had cut power and torn off roof tiles, but was "not too bad".

(67) "But I think to the north - that's the area that has borne the brunt.

(68) Those were the areas worst hit by the earthquake last month."

(69) "In Cebu they have had quite a battering and I spoke to our colleagues and they've had quite strong winds and are locked down in their hotels.

(70) They are waiting to see what the situation is."

(71) Jeff Masters, meteorology director at the private firm Weather Underground, said in a blog post that the damage from Haiyan's winds must have been "perhaps the greatest wind damage any city on Earth has endured from a tropical cyclone in the past century".

(72) Our correspondent says that, while the country is better prepared than for previous storms, it is not clear whether even buildings being used as storm shelters can withstand these winds.

(73) In its path are areas already struggling to recover from a deadly 7.3-magnitude earthquake last month, including the worst-hit island of Bohol where about 5,000 people are still living in tents.

Titel:	Philippine Typhoon Death Toll Feared in Thousands
Thema:	Typhoon Haiyan hits Philippine
Quelle / Datum:	New York Times, 09.11.2013

(74) The powerful typhoon that swept across the Philippines on Friday, one of the strongest storms ever to make landfall, cut a path of destruction through several central islands, leaving the seaside city of Tacloban in ruins and leading to early, unconfirmed estimates of as many as 10,000 dead.

(75) Wire reports quoted the city administrator of Tacloban suggesting that the death toll could reach 10,000 in his city alone.

(76) A police official gave an identical estimate, citing the governor of the area, who had spoken with officials in villages that had been hit, according to wire reports.

(77) The government disaster agency said it could confirm only about 150 deaths so far from Typhoon Haiyan, although the president said he expected the number to rise significantly.

(78) The Red Cross in Manila said earlier on Saturday that its people on the ground were reporting an estimated 1,000 deaths on Leyte Island, where Tacloban is, and about 200 on the neighboring island of Samar.

(79) “The local Red Cross chapter has seen many bodies,” Gwendolyn Pang, the secretary general of the Philippine Red Cross, said in a text message.

(80) “An actual body count has to be done to determine the exact number.”

(81) The destruction, which has taken down phone service in many areas, made confirming any of the accounts difficult.

(82) Some meteorologists said the storm, called Yolanda in the Philippines, hit land with sustained winds above 190 miles per hour, while others reported winds of 150 miles per hour.

(83) On Friday, some in the country thought the Philippines might have been spared high casualties because the storm had moved so quickly, but they did not know that it had caused a serious storm surge, at least in Tacloban.

(84) Photos and television footage showed fierce winds ripping tin roofs off homes and sending waves crashing into wooden buildings that splintered under the force.

(85) Large ships were tossed on shore, and vehicles were shown piled up on top of one another.

(86) Video footage from Tacloban showed ocean water rushing through the streets of the city, which has an estimated population of 220,000.

(87) Speaking to Reuters, the manager of the city’s airport, which is on a strip of land that juts into the sea, estimated that water there rose up to 13 feet.

(88) Reuters also quoted a spokesman for the national disaster agency saying many houses in Tacloban were destroyed.

(89) A bicycle taxi driver who lives near the airport told The Associated Press that he and his family had taken refuge in a parked jeep, which was swept away in the roiling waters.

(90) The man, Sandy Torotoro, said that as the vehicle floated by, many people screamed for help as they were swept away, waving their hands above the water.

(91) “But what can we do?” he said.

(92) “We also needed to be helped.”

(93) The Social Welfare and Development Department said that the storm affected 4.28 million people in about 270 towns and cities spread across 36 provinces in the central Philippines.

(94) President Benigno S. Aquino III said at a news briefing on Saturday evening in Manila that he expected there to be “substantially more” deaths than the government had confirmed.

(95) He arrived Sunday in Tacloban, according to a member of Parliament.

(96) The government has been flying in military cargo planes carrying food, clothing and shelters, but blocked roads have made distribution difficult.

(97) A United Nations disaster assessment team visited the area on Saturday.

(98) “The last time I saw something of this scale was in the aftermath of the Indian Ocean tsunami,” Sebastian Rhodes Stampa, the head of the team, said in a statement, referring to the 2004 tsunami that devastated parts of Indonesia and other countries.

(99) “This is destruction on a massive scale.

(100) There are cars thrown like tumbleweed.”

(101) Richard Gordon, the chairman of the Philippine Red Cross, said in an interview that most of the information about damage and casualties was coming out of Tacloban, where the news media and government officials were concentrated, and that he feared there would be “a lot of dead bodies” inland as well.

(102) He said there were also areas out of contact in northern Cebu and on the island of Panay, as well as parts of Palawan and Mindoro.

(103) Defense Secretary Chuck Hagel directed the United States military’s Pacific Command to provide airborne and maritime search and rescue teams and other help, a Department of Defense statement on Saturday said.

(104) According to the National Disaster Risk Reduction and Management Council, the deadliest storm in Philippine history was Tropical Storm Thelma, which killed more than 5,000 people.

Titel:	"Massive destruction" as typhoon kills at least 1,200 in Philippines, says Red Cross
Thema:	Typhoon Haiyan hits Philippine
Quelle / Datum:	Reuters, 09.11.2013

(105) One of the strongest typhoons ever to make landfall devastated the central Philippines, killing more than 1,000 people in one city alone and 200 in another province, the Red Cross estimated on Saturday, as reports of high casualties began to emerge.

(106) A day after Typhoon Haiyan churned through the Philippine archipelago in a straight line from east to west, rescue teams struggled to reach far-flung regions, hampered by washed out roads, many choked with debris and fallen trees.

(107) The death toll is expected to rise sharply from the fast-moving storm, whose circumference eclipsed the whole country and which late on Saturday was heading for Vietnam.

(108) Among the hardest hit was coastal Tacloban in central Leyte province, where preliminary estimates suggest more than 1,000 people were killed, said Gwendolyn Pang, secretary general of the Philippine Red Cross, as water surges rushed through the city.

(109) "An estimated more than 1,000 bodies were seen floating in Tacloban as reported by our Red Cross teams," she told Reuters. (103) "In Samar, about 200 deaths.

(110) Validation is ongoing."

(111) She expected a more exact number to emerge after a more precise counting of bodies on the ground in those regions.

(112) Witnesses said bodies covered in plastic were lying on the streets.

(113) Television footage shows cars piled atop each other.

(114) "The last time I saw something of this scale was in the aftermath of the Indian Ocean Tsunami," said Sebastian Rhodes Stampa, head of the U.N. Disaster Assessment Coordination Team sent to Tacloban, referring to the 2004 earthquake and tsunami.

(115) "This is destruction on a massive scale.

(116) There are cars thrown like tumbleweed and the streets are strewn with debris."

(117) The category 5 "super typhoon" weakened to a category 4 on Saturday, though forecasters said it could strengthen again over the South China Sea en route to Vietnam.

(118) Authorities in 15 provinces in Vietnam have started to call back boats and prepare for possible landslides.

(119) Nearly 300,000 people were moved to safer areas in two provinces alone - Da Nang and Quang Nam - according to the government's website.

(120) The Philippines has yet to restore communications with officials in Tacloban, a city of about 220,000.

(121) A government official estimated at least 100 were killed and more than 100 wounded, but conceded the toll would likely rise sharply.

(122) The national disaster agency has yet to confirm the toll but broken power poles, trees, bent tin roofs and splintered houses littered the streets of the city about 580 km (360 miles) southeast of Manila.

(123) "IT WAS LIKE A TSUNAMI"

(124) The airport was nearly destroyed as raging seawaters swept through the city, shattering the glass of the airport tower, levelling the terminal and overturning nearby vehicles.

(125) "Almost all houses were destroyed, many are totally damaged.

(126) Only a few are left standing," said Major Rey Balido, a spokesman for the national disaster agency.

(127) Local television network ABS-CBN showed images of looting in one of the city's biggest malls, with residents carting away everything from appliances to suitcases and grocery items.

(128) Airport manager Efren Nagrama, 47, said water levels rose up to four metres (13 ft) in the airport.

(129) "It was like a tsunami.

(130) We escaped through the windows and I held on to a pole for about an hour as rain, seawater and wind swept through the airport.

(131) Some of my staff survived by clinging to trees.

(132) I prayed hard all throughout until the water subsided."

(133) Across the country, about a million people took shelter in 37 provinces after President Benigno Aquino appealed to those in the typhoon's path to leave vulnerable areas.

(134) "For casualties, we think it will be substantially more," Aquino told reporters.

(135) Officials started evacuating residents from low-lying areas, coastlines and hilly villages as early as three days before the typhoon struck on Friday, officials said.

(136) But not all headed the call to evacuate.

(137) "I saw those big waves and immediately told my neighbours to flee," said Floremil Mazo, a villager in southeastern Davao Oriental province.

(138) Meteorologists said the impact may not be as strong as feared because the storm was moving so quickly, reducing the risk of flooding and landslides from torrential rain, the biggest causes of typhoon casualties in the Philippines.

(139) Ferry services and airports in the central Philippines remained closed, hampering aid deliveries to Tacloban, although the military said three C-130 transport planes managed to land at its airport on Saturday.

(140) At least two people were killed on the tourist destination island of Cebu, three in Iloilo province and another three in Coron town in southwestern Palawan province, radio reports said.

(141) "I never thought the winds would be that strong that they could destroy my house," LynLyn Golfan of Cebu said in a television interview while sifting through the debris.

(142) By Saturday afternoon, the typhoon was hovering 765 km west of San Jose in southwestern Occidental Mindoro province, packing winds of a maximum 185 kph, with gusts of up to 220 kph.

(143) The storm lashed the islands of Leyte and Samar with 275-kph wind gusts and 5-6 metre (15-19 ft) waves on Friday before scouring the northern tip of Cebu province.

(144) It weakened slightly as it moved west-northwest near the tourist island of Boracay, later hitting Mindoro island.

(145) Haiyan was the second category 5 typhoon to hit the Philippines this year after Typhoon Usagi in September.

(146) An average of 20 typhoons strike every year, and Haiyan was the 24th so far this year.

(147) Last year, Typhoon Bopha flattened three towns in southern Mindanao, killing 1,100 people and causing damage of more than $1 billion.

Titel:	Over 100 killed in C. Philippine city Tacloban by Typhoon
Thema:	Typhoon Haiyan hits Philippine
Quelle / Datum:	Xinhua, 09.11.2013

(148) Over 100 people were reported killed and 100 others were injured in central Philippine city of Tacloban alone in the aftermath of super typhoon Haiyan (local name Yolanda) , the Civil Aviation Authority of the Philippines (CAAP) said Saturday.

(149) CAAP Director General William Hotchkiss III, quoting reports from the agency's personnel on the ground, said that the presumed dead bodies were lying on the streets near the Tacloban City airport with around 100 more injured and requesting for medical evacuation and additional medical personnel.

(150) He said that the reports were gathered when the team composed of local CAAP members, Aviation Security group and airport personnel started clearing the airport runway since early Saturday morning.

(151) There were no casualties at the airport, but communication was limited due to brownout, Hotchkiss said.

(152) He ordered his deputy for operations Captain John Andrews to fly to Tacloban airport, bringing needed supplies, food, medicine and another set of CAAP communication equipment to the airport.

(153) Airports in Iloilo, Caticlan, Romblon, Dumaguete, Bacolod, Masbate, Legaspi and Surigao are now back to normal operations, while Tacloban and Busuanga were still closed due to severe damage brought by Haiyan, he added.

(154) The state weather forecasting agency said that as of 10:00 a.m. local time, the eye of the typhoon was located based on all available data at 549 km west of San Jose in northern province of Occidental Mindoro.

(155) It has a maximum sustained wind of 175 km per hour near the center and gustiness of up to 210 kph.

(156) Typhoon Haiyan continues to move over the South China Sea and is expected to exit the Philippines Saturday afternoon.

Appendix B. Survey 2 and A Test for Contradiction

Survey 2: Guidelines for Identifying Contradictions and Contrarieties

Introduction

Thank you for participating in the second stage of the survey on contradictions. This survey is a pilot task for development of a system for machine-based contradiction detection. On the next pages you'll find 424 items of two text pieces on one of the nine world events: an armed invasion of Russia in Crimea (February-March 2014), a fire in a Brazil nightclub (January 2013), first protests after the killing of a black teen in Ferguson (August 2014), the mysterious vanishing of the Malaysia Airlines Flight MH370 plane (March 2014), the shooting down of the Malaysia Airlines Flight MH17 (July 2014), the military coup in Thailand (May 2014), the natural catastrophe in the Philippines (November 2013), the crash of the Costa Concordia ship (January 2012) as well as the unexpected death of Whitney Houston (February 2012). The items have been collected manually in the news texts during Survey 1. Some of the text pieces are contradictory and contrary to each other and some are not. We wish to identify the contradictory and contrary ones. For this, please read the text items carefully and make your decision following the test for contradiction below. Finally, select the appropriate box to confirm your decision. For items which are contradictions or contraries select a "YES" box and for non-contradictions and non-contrarieties select a "NO" box. If you are not sure, go for the solution towards which you lean more. You are allowed to use all kinds of supplementary material such as bilingual dictionaries, Wikipedia, etc. for processing the survey.

Preparation

Finding contradictions is a sophisticated task. Especially, if English is not your native language. The task requires a lot of concentration. Contradictions and contrarieties are in most cases implicit and require the ability of logical inferencing. So you will need to use your world knowledge and be able to make inferences in order to recognize contradictions. Don't process the survey if you are tired. It can have an undesirable impact on your decisions. Remember, the results of your participation are very important. They determine how valid and reliable the data for the further research will be. How long it will take you to complete the survey - one or more days - is up to you.

Instruction

Please read carefully two pieces of text. Use a dictionary to translate unknown words if needed. Assure yourself that you understand the content of the text pieces.

Regard the text pieces of an item as related to the same context. Compatible noun phrases (nouns and pronouns - person, place, thing and their modifiers such as articles, possessive nouns and pronouns, adjectives etc.) should be treated as co-referent (=refer to the same place, person or thing) in the absence of clear countervailing evidence. For example, in the sentence pair below you should assume that Houston and the singer refer to the same person. If you still like to check whether the events are co-referent you are welcome to use the attached file *Corpus of Contradictions*. The *n* (id of the sentence in the article) and *article_id* (id of the article in the *Corpus of Contradictions*) of the items show you where the corresponding text piece comes from.

Example of an item:

Authorities have said that police and fire officials were called to Houston's room at the Beverly Hilton Hotel at 3:43 p.m. Saturday after her bodyguard found her unconscious body in a bathtub.

The singer was found unconscious and submerged in the bathtub of her room at the Beverly Hilton by her hairdresser Saturday afternoon, according to TMZ. (ID 201)

Regard the text pieces of an item as stated at the same time. If you are not sure, you are welcome to use the attached file *Corpus of Contradictions*. The time references (*Monday, yesterday, two days ago etc.*) occurring in the text pieces should be assumed to be co-referent (=refer to the same point in time) in the absence of clear countervailing evidence.

Please evaluate the parts of an item as contradictory or contrary **ONLY IF** you are able to state a clear base for a contradiction or a contrariety.

Test for Contradiction