Saarland University's Participation in the Second Shared Task on Source, Subjective Expression and Target Extraction from Political Speeches (STEPS-2016)

Michael Wiegand, Nadisha-Marie Aliman, Tatjana Anikina, Patrick Carroll, Margarita Chikobava, Erik Hahn, Marina Haid, Katja König, Leonie Lapp, Artuur Leeuwenberg, Martin Wolf, Maximilian Wolf
Spoken Language Systems, Saarland University
D-66123, Saarbrücken, Germany
michael.wiegand@lsv.uni-saarland.de

Abstract

We report on the two systems we built for the second run of the shared task on *Source, Subjective Expression and Target Extraction from Political Speeches (STEPS)*. The first system is a rule-based system relying on a predicate lexicon specifying extraction rules for verbs, nouns and adjectives, while the second is a supervised system trained on the adjudicated test data of the previous run of this shared task.

1 Introduction

In this paper, we describe our two systems for the second run of the shared task on *Source, Subjective Expression and Target Extraction from Political Speeches (STEPS)* organized by the Interest Group on German Sentiment Analysis (IGGSA). In that task, both *opinion sources*, i.e. the entities that utter an opinion, and *opinion targets*, i.e. the entities towards which an opinion is directed, are extracted from German sentences. The opinions themselves have also to be detected automatically. The sentences originate from debates of the Swiss Parliament (*Schweizer Bundesversammlung*).

The first system is a rule-based system relying on a predicate lexicon specifying extraction rules for verbs, nouns and adjectives, while the second is a supervised classifier trained on the adjudicated test data of the previous edition of this shared task (Ruppenhofer et al., 2014).

2 Rule-based System

Our rule-based system is an extension of the rule-based system built for the first edition of this shared task as described in Wiegand et al. (2014). The pipeline of the rule-based system is displayed in Figure 1. The major assumption that underlies this system is that the concrete opinion sources and targets are largely determined by the opinion predicate[1] by which they are evoked. Therefore, the task of extracting opinion sources and targets is a lexical problem, and a lexicon for opinion predicates specifying the argument position of sources and targets is required. For instance, in Sentence (1), the sentiment is evoked by the predicate *liebt*, the source is realized by its subject *Peter* while the target is realized by its accusative object *Maria*.

(1) $[\text{Peter}]^{\textit{source}}_{\textit{subj}}$ **liebt** $[\text{Maria}]^{\textit{target}}_{\textit{obja}}$.
(Peter loves Maria.)

With this assumption, we can specify the demands of an opinion source/target extraction system. It should be a tool that given a lexicon with argument information about sources and targets for each opinion predicate

- checks each sentence for the presence of such opinion predicates,
- syntactically analyzes each sentence and
- determines whether constituents fulfilling the respective argument information about sources and targets are present in the sentence.

In the following, we briefly describe the linguistic processing (Section 2.1) and the mechanism for extracting rules (Section 2.2). Then, we introduce the extensions we applied for this year's submission (Section 2.3). For general information regarding the architecture of the system, we refer the reader to Wiegand et al. (2014).

The version of the rule-based system that has been revised for this year's shared task has been made publicly available[2] allowing researchers to

[1] We currently consider opinion verbs, nouns and adjectives as potential opinion predicates.

[2] https://github.com/miwieg/german-opinion-role-extractor

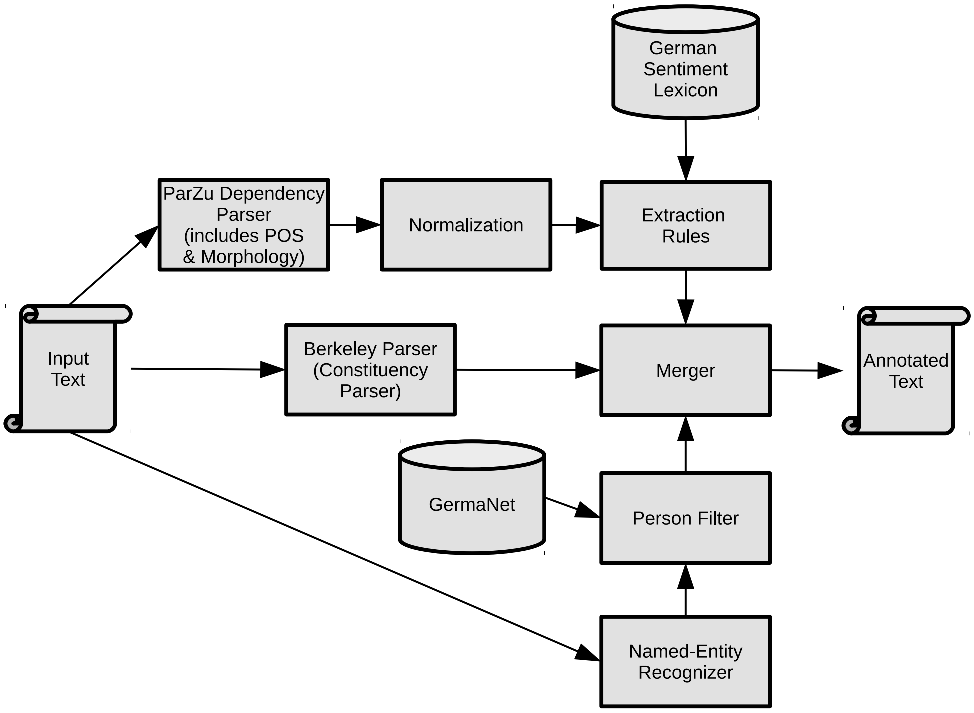


Figure 1: Processing pipeline of the rule-based system.

test different sentiment lexicons with different argument information about opinion sources and targets.

2.1 Linguistic Processing

Even though the data for this task already come in a parsed format, we felt the need to add further linguistic information. In addition to the existing constituency parse provided by the Berkeley parser (Petrov et al., 2006), we also included dependency parse information. With that representation, relationships between opinion predicates and their sources and targets can be formulated more intuitively.[3]

As a dependency parser, we chose *ParZu* (Sennrich et al., 2009). We also carried out some normalization on the parse output in order to have a more compact representation. To a large extent, the type of normalization we carry out is in line with the output of dependency parsers for English, such as the Stanford parser (de Marneffe et al., 2006). It is included since it largely facilitates writing extraction rules. The normalization includes

(a) active-passive normalization

(b) conflating several multi-edge relationships to one-edge relationships

(c) particle-verb reconstruction

These normalization steps are explained in more detail in Wiegand et al. (2014).

We also employed a semantic filter for the detection of opinion sources. Since such entities can only represent persons or groups of persons, we employed a named-entity recognizer (Benikova et al., 2015) to recognize person names and GermaNet (Hamp and Feldweg, 1997) to establish that a common noun represents a person or a group of persons.

2.2 The Extraction Rules

The heart of the rule-based system is a lexicon that specifies the (possible) argument positions of sources and targets. So far, there does not exist any lexicon with that specific information which is why we came up with a set of default rules for the different parts of speech. The set of opinion predicates are the subjective expressions from the PolArt system (Klenner et al., 2009). Every mention of such expressions will be considered as a mention of an opinion predicate, that is, we do not carry out any subjectivity word-sense disambiguation (Akkaya et al., 2009).

[3] As a matter of fact, the most appropriate representation for that task is semantic-role labeling (Ruppenhofer et al., 2008; Kim and Hovy, 2006; Wiegand and Klakow, 2012), however, there currently do not exist any robust tools of that kind for German.

These default extraction rules are designed in such a way that for a large fraction of opinion predicates with the pertaining part of speech they are correct. The rules are illustrated in Table 1. We currently have distinct rules for verbs, nouns and adjectives. All rules have in common that for every opinion predicate mention, at most one source and at most one target is assigned. The rules mostly adhere to the dependency relation labels of ParZu.[4]

The rule for verbs assumes sources in subject and targets in object position (1). Note that for targets, we specify a priority list. That is, the most preferred argument position is a dative object (*objd*), the second most preferred position is an accusative object (*obja*), etc. In computational terms, this means that the classifier checks the entire priority list (from left to right) until a relation has matched in the sentence to be classified. For prepositional complements, we also allow a wildcard symbol (*pobj-**) that matches all prepositional complements irrespective of its particular head, e.g. *über das Freihandelsabkommen* (*pobj-ueber*) in (2).

(2) [Deutschland und die USA]$^{source}_{subj}$ **streiten** [über das Freihandelsabkommen]$^{target}_{pobj-ueber}$.
(Germany and the USA quarrel over the free trade agreement.)

For nouns, we allow determiners (possessives) (3) and genitive modifiers (4) as opinion sources whereas targets are considered to occur as prepositional objects.

(3) [Sein]$^{source}_{det}$ **Hass** [auf die Regierung]$^{target}_{pobj-auf}$ ...
(His hatred towards the government ...)

(4) Die **Haltung** [der Kanzlerin]$^{source}_{gmod}$ [zur Energiewende]$^{target}_{pobj-zu}$ ...
(The chancellor's attitude towards the energy revolution ...)

The rule for adjectives is different from the others since it assumes the source of the adjective to be the speaker of the utterance. Only the target has a surface realization. Either it is an attributive adjective (5) or it is the subject of a predicative adjective (6).

[4]The definition of those dependency labels is available at `https://github.com/rsennrich/ParZu/blob/master/LABELS.md`

Part of Speech	Source	Target
verb	subj	objd, obja, objc, obji, s, objp-*
noun	det, gmod	objp-*
adjective	*author*	attr-rev, subj

Table 1: Extraction rules for verb, noun and adjective opinion predicates.

(5) Das ist ein [guter]$^{target}_{attr-rev}$ **Vorschlag**.
(This is a good proposal.)

(6) [Der Vorschlag]$^{target}_{subj}$ ist **gut**.
(The proposal is good.)

Our rule-based system is designed in such a way that, in principle, it would also allow more than one opinion frame to be evoked by the same opinion predicate. For example, in *Peter überzeugt Maria*/*Peter convinces Maria*, one frame sees *Peter* as source and *Maria* as target, and another frame where the roles are switched. Our default rules do not include such cases, since such property is specific to particular opinion predicates.

2.3 Extensions

In this subsection, we present the extensions we added to the existing rule-based system from the previous iteration of this shared task.

2.3.1 Partial Analysis

Our system has been modified in such a way that it can now accept a partial analysis as input and process it further. By that we mean the existing annotation of subjective expressions as specified by the subtask of this shared task. Given such input, the system just assigns sources and targets for these existing expressions. (We also implemented another mode in which the opinion predicates according the given sentiment lexicon would additionally be recognized including their opinion roles.) Opinion predicates are typically ambiguous; our lexicon-based approach is therefore limited. This is a well-known and well-researched problem. On the other hand, the task of extracting opinion sources and targets given some opinion predicates is a completely different task, which is comparatively less well researched. Our mode allowing partial analysis as input should allow researchers interested in opinion role extraction to have a suitable test bed without caring for the detection of subjectivity.

2.3.2 Grammatically-Induced Sentiment

An important aspect of opinion-role extraction that was completely ignored in the initial version of the rule-based system is the sentiment that is not evoked by common opinion predicates but sentiment that is evoked by certain grammatical constructions. We focus on certain types of modalities (7) and tenses (8). Such type of sentiment is detected without our extraction lexicon (§2.2).

(7) [Kinder **sollten** nicht auf der Straße spielen.]^{target} *source: speaker*
(Children should not play on the street.)

(8) [Er **wird** mal ein guter Lehrer sein.]^{target} *source: speaker*
(He is going to become a good teacher.)

(9) Der Puls des Patienten *wird* gemessen.
(The patient's pulse is measured.)

It is triggered by certain types of lexical units, that is, modal verbs, such as *sollte*, or auxiliary verbs, such as *werden*. However, unlike the lexical units from our extraction lexicon, some of these verbs require some further disambiguation. For instance, the German auxiliary *werden* is not exclusively used to indicate future tense as in (8) but it is also used for building passive voice (9). Therefore, our module carries out some contextual disambiguation of these words.

Grammatically-induced sentiment also systematically differs from lexical sentiment in the way in which opinion roles are expressed. While for lexical sentiment, the argument position of the sources and targets is dependent on the specific lexical unit that conveys the sentiment and therefore has to be specified by lexical rules, the types of grammatically-induced sentiment that we cover share the same argument positions for sources and targets. Typically, the source is the speaker of the utterance and the target is the entire sentence in which the tense or modal occurs. Of course, in case of compound sentences, the scope of the target is only restricted to the clause in which the auxiliary/modal verb occurs (10).

(10) [Er **wird** mal ein guter Lehrer sein]^{target}, da er gut erklären kann. *source: speaker*
(He is going to become a good teacher since he can explain things well.)

2.3.3 Morphological Analysis

Opinion sources are typically persons or groups of persons. In order to ensure that only NPs that match this semantic type are classified as sources, we employed a semantic filter that used the prediction of a named-entity recognizer in case of proper nouns and GermaNet (Hamp and Feldweg, 1997) in case of common nouns. The latter approach, however, is limited considering the high frequency of compounding in German. We observed that in case an opinion source was represented by a compound, such as *SPD-Landtagsabgeordneter*, it could not be established as a person since that term was not included in GermaNet. We examined whether this coverage problem could be solved by morphologically analyzing those compounds and then only looking up their heads (e.g. *Abgeordneter*) which are more likely to be included in GermaNet. A similar experiment was carried out to match opinion predicates in compounds (e.g. *Frühjahrsaufschwung* or *Telefonterror*). Our initial experiments with *morphisto* (Zielinski and Simon, 2009), however, showed no improvement in either opinion source extraction or subjectivity detection due to the high ambiguity in noun compound structures.

2.3.4 Normalizing Conjunctions

In the original version of our system we already incorporated a set of normalization steps of simplifying the dependency parse (Wiegand et al., 2014). The result was a more compact representation of sentences that abstracts from the surface realization of a sentence. This made it simpler to state extraction rules for the extraction of opinion sources and targets. In our submission for this year's task, we added a further normalization step dealing with conjunctions. The original dependency parse typically only directly connects one conjunct with the syntactic roles relevant for opinion roles. For instance, in Figure 2(a) only *lügt* is connected with *Er* by a subject relation. Therefore, our original system would only be able to establish that *Er* is some opinion role of *lügt*. In such cases, we also add another edge with the subject relation connecting the second conjunct (*betrügt*) and its subject (Figure 2(b)).

We also incorporate a set of rules to handle coordination for predicative and attributive adjectives.[5]

[5]For nouns, we could not figure out unambiguous relations where adding further edges would have increased the extraction of sources or targets.

While for predicative adjectives, the subjective relation has to be duplicated, for attributive adjectives, the edge *attr* needs to be duplicated (see Figure 3).

2.3.5 Alternative NLP Tools

We critically assessed the choice of NLP tools used in our original system and compared them with alternatives.

As far as both constituency and dependency parsing is concerned, it was not possible for us to find more effective alternatives. Either the corresponding parser could not be converted in our existing format so that the system would still work as before, or, the parsing output was notably inferior and produced worse extraction performance when incorporated in our processing pipeline.

As far as named-entity recognition is concerned, we replaced the original tagger, the German model from the Stanford named-entity tagger (Faruqui and Padó, 2010), by a more recent tagger, i.e. GermaNER (Benikova et al., 2015). We primarily decided in favour of this replacement since the Stanford named-entity tagger occasionally overrides the given sentence-boundary detection.

2.4 Multiword Expressions

Not every opinion predicate is a unigram token. So far, the only *multiword expressions* conveying sentiment that our system was able to process were phrasal verbs, as *gab auf* in (11).

(11) Er **gab** das Rauchen vor 10 Jahren **auf**.
(He gave up smoking 10 years ago.)

We modified our system in such a way that extraction rules can now also be specified for arbitrary multiword expressions. Matching multiword expressions in German sentences is not trivial since

- multiword expressions can be discontinuous sequences of tokens (e.g. (12), (13)),
- the order of tokens between the canonical form and mentions in specific sentences may vary (e.g. (13)), and
- several tokens between the canonical form and mentions in specific sentences may differ (e.g. reflexive pronoun *sich* in (12)).

In order to account for these properties, our matching algorithm considers the dependency parse of a sentence. We identify a multiword expression if the tokens of a particular expression are all directly connected via edges in a dependency parse. The multiword expressions must hence form a connected subgraph of the parse in which all tokens of the multiword expression and only those are included.

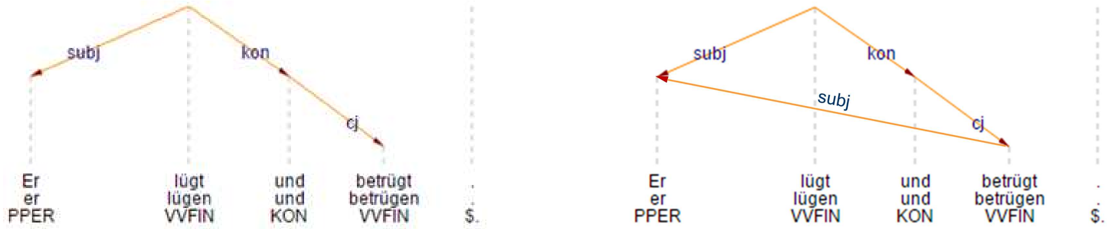
(12) *sich benehmen wie die Axt im Walde* (*act like a brute*):
Er sagte, dass ich **mich benehmen** würde, **wie die Axt im Walde**.
(He told me that I acted like a brute.)

(13) *sein wahres Gesicht zeigen* (*to show one's true colours*):
Unter Alkoholeinfluss **zeigte** er **sein wahres Gesicht**.
(Under the influence of alcohol, he showed his true colours.)

Since we are not aware of any publicly available lexicon for multiword expressions, we extract them automatically from a German corpus. For that, we use the parsed *deWaC* (Baroni et al., 2009). We focus on those multiword expressions that follow a systematic pattern. We chose *reflexive verbs* (e.g. *sich freuen*, *sich schämen*, *sich fürchten*) and *light-verb constructions* (e.g. *Angst haben*, *Kummer machen*, *Acht geben*). In order to extract reflexive verbs, we extracted opinion verbs frequently occurring with a reflexive pronoun (we restrict the pronoun to be the accusative object of the verb). In order to extract light-verb constructions, we first manually selected a set of common light verbs (e.g. *haben*, *machen*, *geben*) and then looked for opinion nouns that often co-occur with these light verbs (we restrict the opinion noun to be the accusative object of the light verb). In total, we thus extracted about 4700 multiword expressions.

3 Supervised System

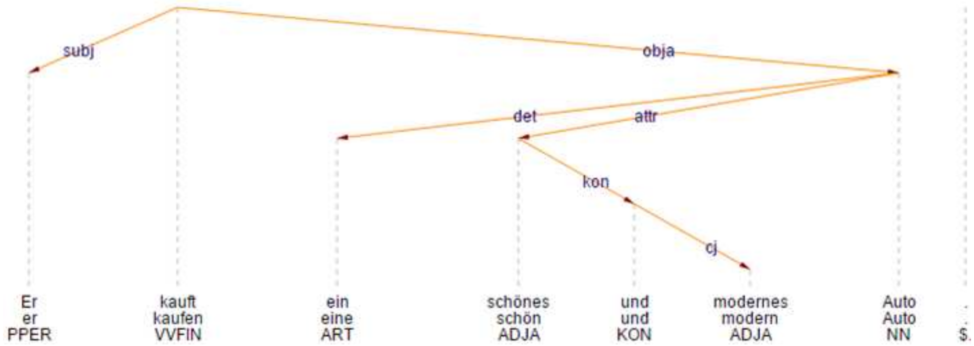
Since we participated in the second edition of this shared task, it meant that we were able to exploit the manually-labeled test data of the previous shared task (Ruppenhofer et al., 2014) as training data for a supervised classifier. In the previous edition, also a supervised system was submitted, however, it only considered as labeled training data texts automatically translated from English to German. Moreover, only opinion sources were con-



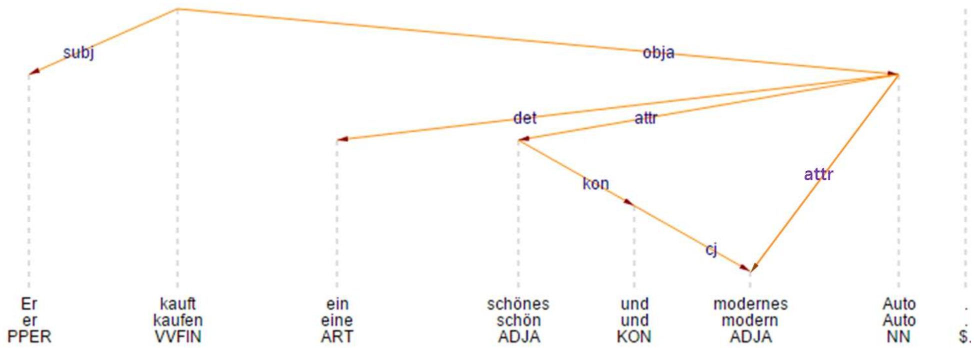
(a) original dependency parse

(b) normalized dependency parse

Figure 2: Illustration of normalizing dependency parses with verb coordination.



(a) original dependency parse



(b) normalized dependency parse

Figure 3: Illustration of normalizing dependency parses with adjective coordination.

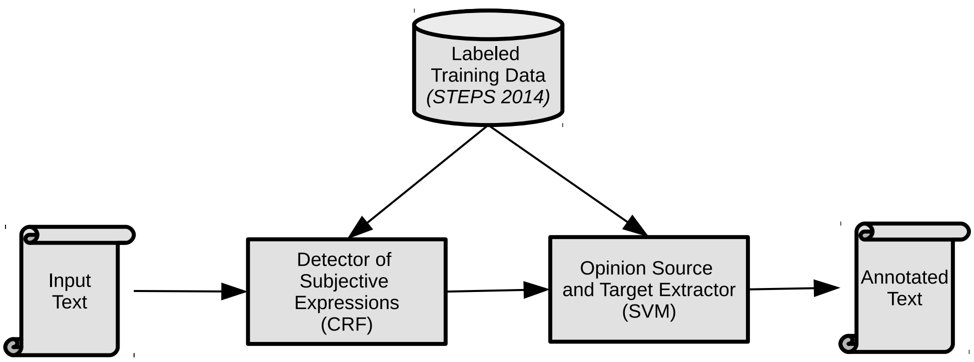


Figure 4: Processing pipeline of the supervised system.

Type	Feature Templates
words	unigram features: target word and its two predecessors/successors bigrams features: bigrams of neighboring words from unigram features
part of speech	unigram features: part-of-speech tag of target word and its two predecessors/successors bigram features: bigrams of neighboring part-of-speech tags from unigram features bigram features: trigrams of neighboring part-of-speech tags from unigram features
sentiment lexicon	is either of the words (window is that of the unigram features) an opinion predicate according to sentiment lexicon

Table 2: Feature templates employed for the CRF classifier to detect subjective expressions.

sidered. We believe that considering actual German text presents a much higher quality of training data than text that has automatically been translated into German.

The processing pipeline of our supervised system is illustrated in Figure 4. For this approach, we employed the same NLP tools as in our rule-based system in order to ensure comparability.

Our supervised system comprises two classifiers: The first is to detect opinion predicates. For that, we employ a conditional random field (Lafferty et al., 2001). As an implementation, we chose *CRF++*[6]. As a motivation, we chose a sequence-labeling algorithm because the task of detecting opinion predicates is similar to other tagging problems, such as part-of-speech tagging or named-entity recognition. The feature templates for our sentiment tagger are displayed in Table 2. We use CRF++ in its standard configuration; as a labeling scheme, we used the simple IO-notation.

The second classifier extracts for an opinion predicate detected by the CRF the corresponding opinion source or target, if they exist. For this support vector machines (SVM) were chosen. As an implementation, we used *SVM^{light}* (Joachims, 1999). The instance space is a set of tuples comprising candidate opinion roles and opinion predicates (detected by the previous sentiment detection). We use different sets of candidate phrases for opinion sources and opinion targets. For opinion sources, the set of candidates is the set of noun phrases in a sentence. Opinion sources are typically persons or groups of persons and, therefore, only noun phrases are eligible to represent such opinion role. Opinion targets, on the other hand, cannot be reduced to one semantic type. Targets can be various types of entities, both animate and inanimate. They can even represent entire propositions. As a consequence, we consider every constituent phrase of a sentence as a candidate opinion target.

[6]`https://code.google.com/p/crfpp/`

SVM were chosen as a learning method since this task deals with a more complex instance space, and SVM, unlike sequence labelers, allow a fairly straightforward encoding of that instance space. The features we employed for this classifier are illustrated in Table 3.

4 Experiments

In this section, we evaluate the 7 runs officially submitted to the shared task. Table 4 displays the different properties of the different runs. The first 5 runs are rule-based systems, while the last run is a supervised system. *Rule-Based-2014* is the best rule-based system run in the previous iteration of this shared task (Wiegand et al., 2014) using the PolArt-sentiment lexicon (Klenner et al., 2009). *Rule-Based-2016-plain* is as *Rule-Based-2014* with various bugs removed. *Rule-Based-2016-gram* is as *Rule-Based-2016* with the module on grammatically-induced sentiment analysis (Section 2.3.2) switched on. *Rule-Based-2016-conj* is as *Rule-Based-2016-gram* but also with normalization of conjunctions (Section 2.3.4) switched on. The last system, *Supervised* is the supervised classifier presented in Section 3.

Table 5 displays the (micro-average) performance (exact matches) of the different configurations on the full task. **SE** evaluates the detection of subjective expressions, **Source** the detection of opinion sources and **Target** the detection of opinion targets.

Table 5 shows that the extensions made to the 2014-system result in some improvement. This improvement is caused by a notable rise in recall. The normalization of conjunctions and the treatment of multiword expressions only produce mild performance increases. We assume that this is due to the fact that, in the test data, there are only few cases of the conjunctions we deal with and also only few cases of the multiword expressions we extracted from a corpus. If one compares the rule-based systems with the supervised

Type	Features
candidate opinion role	phrase label of candidate opinion role (e.g. *NP*, *VP*, *SBAR* etc.)
	lemma of head of phrase representing candidate opinion role
	part-of-speech of head of phrase representing candidate opinion role
	is head of phrase representing candidate opinion role some named entity?
	is candidate opinion role at the beginning of the sentence?
opinion predicate	lemma of opinion predicate
	part-of-speech of opinion predicate
relational	distance between opinion role candidate and opinion predicate
	dependency path from opinion role candidate to opinion predicate
	part-of-speech label of head of opinion role candidate and opinion predicate
	phrase label of opinion role candidate and part-of-speech tag of head of opinion predicate

Table 3: Features employed for the SVM classifier to extract opinion sources and targets.

Run	Properties
Rule-Based-2014	previous system (Wiegand et al., 2014) as it was publicly available
Rule-Based-2016-plain	as Rule-Based-2014 with various bugs removed
Rule-Based-2016-gram	Rule-Based-2016 with module on grammatically-induced sentiment analysis (Section 2.3.2) switched on
Rule-Based-2016-conj	as Rule-Based-2016-gram but also with normalization of conjunctions (Section 2.3.4) switched on
Rule-Based-2016-mwe	as Rule-Based-2016-conj but also with additional multiword expressions as part of the sentiment lexicon (Section 2.4)
Supervised	supervised learning system as discussed in Section 3

Table 4: The different properties of the different runs.

Run	Measure	SE	Source	Target
Rule-Based-2014	Prec	57.01	44.85	47.64
	Rec	15.14	9.70	11.41
	F	23.93	15.50	18.41
Rule-Based-2016-plain	Prec	55.54	41.83	45.27
	Rec	19.90	13.83	12.71
	F	29.30	20.79	19.85
Rule-Based-2016-gram	Prec	56.36	42.04	44.83
	Rec	24.95	18.62	17.63
	F	34.59	25.81	25.30
Rule-Based-2016-conj	Prec	56.36	42.11	45.01
	Rec	24.95	18.67	17.85
	F	34.59	25.88	25.57
Rule-Based-2016-mwe	Prec	57.20	45.52	44.03
	Rec	25.32	18.95	18.53
	F	35.10	26.22	**26.08**
Supervised	Prec	65.42	50.24	32.31
	Rec	41.41	23.25	17.29
	F	**50.72**	**31.79**	22.52

Table 5: Evaluation of the different runs of the Main Task.

Run	Task	Measure	SE	Source	Target
Rule-Based-2016-mwe	Full Task	Prec	57.20	45.52	44.03
		Rec	25.32	18.95	18.53
		F	35.10	26.22	**26.08**
	Subtask	Prec	100.0	59.86	69.24
		Rec	100.0	28.60	28.87
		F	100.0	38.70	**40.75**
Supervised	Full Task	Prec	65.42	50.24	32.31
		Rec	41.41	23.25	17.29
		F	50.72	31.79	22.52
	Subtask	Prec	100.0	59.40	42.60
		Rec	100.0	38.29	31.69
		F	100.0	**46.57**	36.35

Table 6: Evaluation of the Subtask.

system, one notices that on the detection of subjective expressions, the supervised system largely outperforms the rule-based system. Both precision and recall are improved. Obviously, the supervised system is the only classifier capable of disambiguating subjective expressions (Akkaya et al., 2009). Moreover, it seems to detect more subjective expressions than are contained in a common sentiment lexicon (which is the backbone of the subjectivity detection of the rule-based system). The supervised system, however, is less effective on the extraction on targets. Targets are more difficult to extract than sources in general since they can be more heterogeneous linguistic entities. Sources, for instance, are typically realized as noun phrases, whereas targets can be different types of phrases. We also observed that due to the fact that many targets are comparably large spans, parsing errors also affect this type of opinion roles more frequently. On the other hand, the constituents typically representing sources, i.e. (small) noun phrases, can be correctly recognized more easily. The supervised system may also outperform the rule-based system on the extraction of sources, since it can memorize certain entities with a high prior likelihood to be sources. For instance, first person pronouns (e.g. *I* or *we*) are very likely candidates for sources. This type of information cannot be incorporated in the rule-based classifier.

Table 6 compares the best rule-based system (i.e. *System-2016-mwe*) and the supervised system on both the full task and the subtask (again: micro-average performance – exact matches). In the subtask, subjective expressions are already given and only sources and targets have to be extracted. Obviously the subtask is easier which can be seen by the notably higher performance scores on both source and target extraction for both approaches. As on the full task, on the extraction of sources the supervised system outperforms the rule-based system, while on the extraction of targets, the rule-based system outperforms the supervised system.

5 Conclusion

We reported on the two systems we devised for the second edition of the shared task on *Source, Subjective Expression and Target Extraction from Political Speeches (STEPS)*. The first system is a rule-based system relying on a predicate lexicon specifying extraction rules for verbs, nouns and adjectives, while the second is a supervised classifier trained on the adjudicated test data of the previous edition of this shared task.

The supervised classifier scores well on the detection of subjective expressions and opinion sources. The rule-based system produces the best scores for the extraction of targets. Given the general low performance scores, we assume that the task of opinion source and target extraction still requires some further research.

Acknowledgements

This work was partially supported by the German Research Foundation (DFG) under grant WI 4204/2-1.

References

Cem Akkaya, Janyce Wiebe, and Rada Mihalcea. 2009. Subjectivity Word Sense Disambiguation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 190–199, Singapore.

Marco Baroni, Silvia Bernardini, Adriano Ferraresi, and Eros Zanchetti. 2009. The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora. *Language Resources and Evaluation*, 43(3):209–226.

Darina Benikova, Seid Muhie Yimam, Prabhakaran Santhanam, and Chris Biemann. 2015. GermaNER: Free Open German Named Entity Recognition Tool. In *Proceedings of the International Conference of the German Society for Computational Linguistics and Language Technology (GSCL)*, pages 31–38, Essen, Germany.

Marie-Catherine de Marneffe, Bill MacCartney, and Christopher D. Manning. 2006. Generating typed dependency parses from phrase structure parses. In *Proceedings of the Conference on Language Resources and Evaluation (LREC)*, pages 449–454, Genoa, Italy.

Manaal Faruqui and Sebastian Padó. 2010. Training and Evaluating a German Named Entity Recognizer with Semantic Generalization. In *Proceedings of KONVENS*, Saarbrücken, Germany.

Birgit Hamp and Helmut Feldweg. 1997. GermaNet - a Lexical-Semantic Net for German. In *Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, pages 9–15, Madrid, Spain.

Thorsten Joachims. 1999. Making Large-Scale SVM Learning Practical. In B. Schölkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*, pages 169–184. MIT Press.

Soo-Min Kim and Eduard Hovy. 2006. Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text. In *Proceedings of the ACL Workshop on Sentiment and Subjectivity in Text*, pages 1–8, Sydney, Australia.

Manfred Klenner, Angela Fahrni, and Stefanos Petrakis. 2009. PolArt: A Robust Tool for Sentiment Analysis. In *Proceedings of the Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 235–238, Odense, Denmark.

John Lafferty, Andrew McCallum, and Fernando Pereira. 2001. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proceedings of the International Conference on Machine Learning (ICML)*, Williamstown, MA, USA.

Slav Petrov, Leon Barrett, Romain Thibaux, and Dan Klein. 2006. Learning Accurate, Compact, and Interpretable Tree Annotation. In *Proceedings of the International Conference on Computational Linguistics and Annual Meeting of the Association for Computational Linguistics (COLING/ACL)*, pages 433–440, Sydney, Australia.

Josef Ruppenhofer, Swapna Somasundaran, and Janyce Wiebe. 2008. Finding the Source and Targets of Subjective Expressions. In *Proceedings of the Conference on Language Resources and Evaluation (LREC)*, pages 2781–2788, Marrakech, Morocco.

Josef Ruppenhofer, Roman Klinger, Julia Maria Struß, Jonathan Sonntag, and Michael Wiegand. 2014. IGGSA Shared Tasks on German Sentiment Analysis (GESTALT). In G. Faaß and J. Ruppenhofer, editors, *Workshop Proceedings of the KONVENS Conference*, pages 164–173, Hildesheim, Germany. Universität Hildesheim.

Rico Sennrich, Gerold Schneider, Martin Volk, and Martin Warin. 2009. A New Hybrid Dependency Parser for German. In *Proceedings of the German Society for Computational Linguistics and Language Technology (GSCL)*, pages 115–124, Potsdam, German.

Michael Wiegand and Dietrich Klakow. 2012. Generalization Methods for In-Domain and Cross-Domain Opinion Holder Extraction. In *Proceedings of the Conference on European Chapter of the Association for Computational Linguistics (EACL)*, pages 325–335, Avignon, France.

Michael Wiegand, Christine Bocionek, Andreas Conrad, Julia Dembowski, Jörn Giesen, Gregor Linn, and Lennart Schmeling. 2014. Saarland University's Participation in the GErman SenTiment AnaLysis shared Task (GESTALT). In G. Faaß and J. Ruppenhofer, editors, *Workshop Proceedings of the KONVENS Conference*, pages 174–184, Hildesheim, Germany. Universität Hildesheim.

Andrea Zielinski and Christian Simon. 2009. Morphisto – An Open Source Morphological Analyzer for German. In *Proceedings of the 2009 Conference on Finite-State Methods and Natural Language Processing: Post-proceedings of the 7th International Workshop FSMNLP 2008*, pages 224–231. IOS Press Amsterdam, The Netherlands.