Speaker Attribution in Cabinet Protocols

Josef Ruppenhofer, Caroline Sporleder, Fabian Shirokov

Department of Computational Linguistics, Saarland University, Saarbrücken
Building C 74, PO Box Postfach 15 11 50
{josefr, csporled, fabians}@coli.uni-saarland.de

Abstract

Historical cabinet protocols are a useful resource which enable historians to identify the opinions expressed by politicians on different subjects and at different points of time. While cabinet protocols are often available in digitized form, so far the only method to access their information content is by keyword-based search, which often returns sub-optimal results. We present a method for enriching German cabinet protocols with information about the originators of statements. This requires automatic speaker attribution. In order to avoid costly manual annotation of training data, we design a rule-based system which exploits morpho-syntactic cues. Unlike many other approaches, our method can also deal with cases in which the speaker is not explicitly identified in the sentence itself. This is an important capability as 45% of all sentences in the data constitute reported speech whose speakers are not explicitly marked. Our system is able to detect implicit speakers by taking into account signals of speaker continuity. We show that such a system obtains good results, especially with respect to recall which is particularly important for information access.

1. Introduction

A growing number of projects is aimed at making cultural heritage data more accessible by digitizing them and making them publically searchable. Ultimately one would like to do more than simple keyword-based search. Support for more sophisticated search typically requires enriching the data with additional information such as semantic disambiguation, explicit structuring, or introducing hyper-links between documents.

In this paper, we address the enrichment task of attributing speech events in cabinet protocols to their speakers. Such information allows historians to search systematically for statements made by a particular politician. This capability is important because statements frequently reflect opinions of their speakers and can also provide information about which facts where known by a particular person at a given time. Such information is often crucial for historical research.

Speaker attribution in cabinet protocols is not a trivial task since they typically do not contain direct speech but summaries in indirect speech. This means that some kind of deeper analysis is required. Speaker attribution is essentially a classification task and can be modelled by supervised machine learning. However, this requires manually labelled training data, which are typically not available for the cultural heritage domains and too expensive to create from scratch. We investigate an alternative approach which exploits the fact that there are certain linguistic regularities in this text type that can be utilized by a rule-based system. One important linguistic cue is the use of the subjunctive. For instance in (1), the use of subjunctive mood and present tense on the main verb mark sentence (2) as a continuation of the chancellor's statement in the preceding sentence. (A formulation using indicative mood would be understood as expressing the text writer's opinion.) In our meeting protocols, such morpho-syntactic cues are used with great regularity and we show that by relying on such cues it is possible to identify the speakers of sentences without explicit attribution. Such sentences are not rare: in data we annotated for agreement testing, 45% of all sentences continued

reported speech begun in a previous sentence.

(1) a. Der Bundeskanzler **erklärt**, daß er dem Kabinett zur Saarfrage alles gesagt habe, was er wisse.

   b. *The chancellor states that he has told the cabinet everyting about the Saar question that he knows.*

(2) a. Seitdem SEI nichts geschehen und es werde auch nichts geschehen.

   b. *Since then nothing had happened and nothing would happen.*

We present a fairly shallow, rule-based approach to speaker attribution in cabinet protocols. Before describing our approach in detail, we give an overview of related work (Section 2.) and of the data and annotation scheme we used (Section 3.).

2. Related Work

An important theoretical work on the issue of speaker attribution is Bergler's thesis (1992). Wiebe's (1990) dissertation has a broader focus of psychological subjects and provides a first implemented system for tracking point of view in narratives.

A lot of closely related applied work occurs in the area of sentiment analysis, where automatic systems are built to identify expressions of opinion. Recent systems have also begun tackling the task of finding the topics of opinions. Only very few try to find the sources of opinions. Opinion-related work is different from ours in that we are interested in statements of fact as much as in statements of opinion. Further, while certain applications like review mining e.g. Blair-Goldensohn et al. (2008) can simply assume a single speaker/opinion holder for each text, the speech events in our data have a variety of speakers. Another important distinction between our work and opinion holder extraction is that we also identify the speakers of non-opinionated sentences. In addition, a key component of our algorithm is

that by looking for positive cues for speaker continuity, we are also able to identify speakers for sentences where no communicative, evaluative or cognitive predicate is present. By contrast, Bethard et al. (2004), look only for opinion holders occurring as arguments of propositional attitude verbs. Kim and Hovy (2006) rely on the roles by a role labeler as candidates for the status of opinion topic or holder, and Choi et al. (2005) also require that "the source phrases should be directly related to an opinion expression". Only Seki et al.'s (2009) system for opinion holder identification uses information from prior sentences, opinionated or not, as evidence for the identification of speakers in the current opinionated sentence. Finally, the work by Krestel et al. (2008) on automatically finding sources of reported speech is also closely related. However, it focuses on English language news and, like sentiment research, it only targets reported speech with an explicit reporting clause, while our system can also deal with so-called 'free indirect speech'.

3. Data and Annotation

Our data are taken from the minutes of the weekly meetings of the German cabinet during the period 1949-1960. The minutes are available to the public on request from the Bundesarchiv (German federal archive).[1] We automatically identified sentence boundaries in the documents by applying the OpenNLP[2] sentence splitter. From our collection of 58,310 sentences, we then randomly extracted two data sets: a development set of 687 sentences and a test set of 400 sentences. The first set was annotated by one of the authors according to pre-defined guidelines outlined below, the second set was annotated by the remaining two authors. We had to discard some examples due to sentence splitting mistakes. Our final data sets contained 566 sentences (development) and 323 sentences (test). To assess the reliability of the annotation, we computed inter-annotator agreement on the test set The average inter-annotator f-score achieved was .87 on a strict measure and .88 on a looser one (see Section 5. for the definitions of the measures).

The annotation task consisted of recording for every sentence the set of speakers for all actual present or past speech events and private states (Wiebe et al. 2005) expressed in the sentence, as shown in (3).

(3)   a.   <sentence id="149" hasSpeaker="281,5"> <person id="281"> Der Bundesinnenminister </person>**schließt** sich der **Auffassung**<person id="5"> des Bundeskanzlers </person>an, wird den Entwurf noch zurückhalten und verschiebt die von ihm vorgesehenen Besprechungen. </sentence>

        b.   *The Secretary of the Interior concurs with the opinion of the Chancellor, is going to hold back the proposal for a while, and postpones the talks he had planned.*

Note that in the annotation process, references to offices are resolved to their holders. Thus, for (3) we record id 281

(Robert Lehr) as speaker rather than using an id for the office of Secretary of the Interior. The set of persons available for annotation as speakers was limited to an almost exhaustive biographical database from the Bundesarchiv, which contains the most important politicians and officials from the period. In addition, a virtual speaker was created to represent the cabinet as a whole. The total number of potential speakers available for annotation is 1932.

Whereas newspaper texts usually include statements by the writer as well as quoted sources, almost all sentences in the minutes report utterances by the meeting participants. Only a few sentences contain background or meta information given by the minute takers. Since these sentences are not of interest to our application scenario, they were assigned the speaker value "unknown". This value was also assigned when the annotators could not identify a speaker at all or where the speaker was an identifiable politician or organization which, however, lacked an entry in the database.

In line with Wiebe et al.'s practice of marking future or hypothetical speech events and private states as insubstantial (Wiebe et al. (2005)), we leave these cases unannotated. For instance, in (4), while we make a record of the consensus reached, we do not capture the speech event of publication that the cabinet decides to refrain from.

(4)   a.   Der Bundesminister für Angelegenheiten der Vertriebenen trägt das Memorandum der Alliierten Hohen Kommission vor . Es besteht **Übereinstimmung** , daß dieses der Öffentlichkeit nicht bekanntzugeben ist .

        b.   *The Secretary for the Affairs of Displace People presents the memorandum by the Allied High Commission. There is consensus that it will not be made known to the public.*

Not annotating the insubstantial cases, while making the annotation task easier, is somewhat detrimental to our application scenario since it is often important to know when somebody chooses not to supply or volunteer information that they have.

Note that our annotation guidelines allow for a sentence to be associated with more than one speaker. For instance, in (3), the Secretary of the Interior is reported by the minutes as concurring with the Chancellor. However, the embedding of the two speech events is not represented: both speakers simply appear on the list of speakers associated with the sentence. While one would ideally like to represent the hierarchical embedding of opinions, we consider our coarser results to be acceptable for the kinds of information users that we envision. Historians will want to check the details of who spoke on which topic for themselves rather than relying on summary information provided by an automatic system.

To give readers a better sense of our data, we present some basic statistics calculated from our agreement data. Table 1 shows that the average sentence contains 1.6 speech events or private states. The average sentence also contains speech acts or private states by more than one speaker. About 17% (84/493) of the speech events or private states are insubstantial, that is, they are future, counterfactual or hypothet-

---

ical. Finally, 14% (58/405) of the speakers associated with private states and speech events are unknown.

	Total	Avg. per S
private states/speech	493	1.6
insubstantial events	84	0.3
speakers	405	1.4
unknown speakers	58	0.2

Table 1: Basic statistics on the test set

Finally note that the 405 speakers associated with the different private states and speech events in our agreement data set corespond to 75 unique individuals. 5 out of the 75 individuals account for 64% for all identifiable speakers: the Cabinet, Adenauer (Chancellor), Schäfer (Secretary of the Treasury), Erhard (Secretrary of Commerce), and Blücher (Secretary for the Marshal Plan). This skew is to some extent a historical artifact: Adenauer ruled throughout the whole period for which we have data and his Cabinet members served long tenures.

4. Linguistic background

The main linguistic cue that we exploit in our work is the highly regular use of the subjunctive mood in formal German in contexts of reported speech. It is exemplified by the passage in (5-9). New speakers typically appear as the subjects of reporting verbs in the indicative mood (5,7, 8). Reporting verbs appear in bold in the examples. The contents of reported speech are typically presented in the subjuntive mood, which is rendered by underlining. Subjunctive mood is used for reported content whether that content appears with a reporting clause, as in (5) or without one, as in (9). Sentence (9) shows the simple case where the use of the subjunctive mood marks a sentence as continuing reported speech from a previous sentence, in this case sentence (8). Sentence (6) represents the interesting case where the private state of one source (the Secretary of Transportation) is embedded under the continuing speech event of another (the undersecretary of State) from a prior sentence, in this case (5).

(5) a. Staatssekretär Hartmann **bemerkt** ergänzend, daß über die in dieser Vorlage angeschnittenen Fragen soeben eine Chefbesprechung 26 stattgefunden <u>habe</u>, die zu keiner Einigung geführt <u>habe</u>.

  b. *Undersecretary of state Hartmann observes in addition that, concerning the issues broached in this proposal, a principals' meeting had taken place just now , which had not produced an agreement.*

(6) a. Überdies <u>**wolle**</u> der Verkehrsminister das Ermäßigungsprogramm umarbeiten und auf Kinder bis zu 25 Jahren ausdehnen.

  b. *On top of that, the secretary of transportation wanted to revise the discount program and extend it to children up to 25 years.*

(7) a. Der Bundesminister für Verkehr **erklärt** hierzu, daß er diese Absicht nicht mehr <u>habe</u>.

  b. *The transportation secretary explains that he no longer has this intention.*

(8) a. Der Bundesminister für Familienfragen **betont** demgegenüber, daß man sich in der genannten Chefbesprechung einig geworden <u>sei</u>.

  b. *The Secretary for Family Affairs stresses, by contrast, that there had been an agreement in the aforementioned principals' meeting.*

(9) a. Man <u>solle</u> vorläufig an der Vorlage festhalten und sie möglicherweise später verbessern.

  b. *One should hold fast to the proposal and improve it later , if possible.*

The passage above also illustrates another important pattern found in the protocols: whenever a potential speaker appears as subject of a sentence, he is typically an actual speaker (at some depth of embedding). This is true for all sentences except (9).

It should be noted that other linguistic cues can also be exploited in analyzing comparable data from languages which may lack the clear morphological cues that German offers. Consider passage (10-13), which is taken from the minutes of a meeting of the British Cabinet on 28 January 1919.[3]

(10) Sir Eric Geddes **said** that it was proposed so to throw the net as to get more men than we require.

(11) The A.S.C on the lines of communication contained a large proportion of the older men.

(12) In the combatant services there were many older men who were `pivotal` N.C.O.'s and who `must` be retained.

(13) He `therefore` did not **see** why it should be necessary to discriminate against the A.S.C.

Sentence (10) clearly introduces reported speech. Sentence (11) might be a statement by the minute taker, based on the words and constructions it contains. However, sentence (12) contains uses of the appraisal *pivotal* and of (epistemic) *must*, which it is very likely the minute taker would not use in a background informational statement. Sentence (13) then confirms the impression that sentences (11) and (12) also represent the views of Geddes by anaphorically referring to them as the reasons for Geddes' opinion stated in the current sentence.

While the cues used in this analysis of the above passage are not a hundred percent reliable individually, we think that in combination they can be useful in the task of attributing content to speakers.

Finally, we note that while the notion of speaker continuity is very useful in analyzing protocols, it may not be as relevant in other genres. An illustration of this is a random

---

[3]The minutes are made available by the British National Archives: http://www.nationalarchives.gov.uk/cabinetpapers/ .

sample of 10 Associated Press newswire stories totalling 4122 words that we collected from the year 2003. This sample contains 122 expressions of speech events and private states. We found that the only type of speaker continuity that occurs is of the type exemplified in (14), where direct speech is continued and where that continuation can be readily recognized from the quotation marks.

(14)   "The domestic leisure market is growing rapidly and now represents over 60 percent of all passengers," Qantas Chief Executive Officer Geoff Dixon said Monday. "Jetstar will concentrate on growing this market with value fares while opening up new destinations."

There were, however, no cases where indirect speech is continued past a reported speech-sentence marked by a reporting verb. This confirms Bergler's (1995) finding that so-called free indirect speech is virtually absent from North American newspaper writing.

5.   Experiments

We use precision (prec), recall (rec) and f-score (f) as our evaluation metrics. **Loose precision** counts a sentence as correctly labeled if at least one of the recognized speakers is correct. **Strict precision** requires all recognized speakers to be correct. Similarly, for **loose recall**, we count a sentence as correctly labeled if at least one of the speakers in the sentence was found by our system. For **strict recall**, we require all of them to have been found.

While we would have liked to use a baseline algorithm used in other work, we decided to construct one of our own because our data contains many instances where (some part of) the content of a sentence needs to be attributed to a speaker that is not explicitly mentioned in that sentence, as seen in (2, 6, 9). As noted in section 1., in the 300-sentence sample that we annotated as agreement data, there were 135 sentences (45%) that continued the speech of a speaker mentioned explicitly only in a preceding sentence. Tools like Krestel et al.'s (2008) Reported Speech Tagger[4] do not look for cases where speakers neeed to be carried over from prior discourse. In addition, Krestel et al.'s tool does not include a grammar for German.

The baseline algorithm we use is centrally based on the intuition that whenever the main clause of the current sentence s_i is headed by a subjunctive mood verb, we can assume speaker continuity with the previous sentence s_{i-1}. Otherwise, we need to associate the speaker of s_i with some referent mentioned in s_i.

Additionally, we look for sentence-initial occurrences of the third person masculine singular pronoun *er*, which typically occurs in cases of same-speaker continuity (of the type "He added that ...").[5] If one of these two rules matches and there exists a prior sentence in the protocol with a known speaker, we carry over that speaker. If there is no

prior sentence s_{i-1} or it lacks a known speaker, we set speaker to the value "unknown". If the first two rules do not fire, we assign as speaker the first person mentioned in sentence s_i, with person mentions being supplied by our Named Entity recognition system. If there are no such mentions, we assign "unknown". To obtain named entity and co-reference information we use a custom-built system that was developed as part of a student project at Saarland University (Atef et al., September 2009). Using a custom-built system proved necessary as most references to people are made not by name, but by the office they hold. Our system can resolve these mentions to specific individuals by connecting document metadata to a database with information on the tenure of the various officials. The system was evaluated on a subset of the data and obtained 79.9% precision, 74.2% recall and 76.9% f-score.

An example where the subjunctive mood rule correctly applies is sentence (9), whose main verb is the subjunctive form *solle*. Sentence (9) continues the reported speech begun in sentence (8) by the Secretary for Family Affairs. In sentences (5,7,8) neither the subjuntive mood nor the *er*-rule fires and therefore the first-encountered potential speaker is assigned speaker status, which is correct in all three cases. Sentence (6) is assigned the speaker from the previous sentence because its main verb is in the subjunctive mood. While this is correct, the system fails to also record the Secretary of Transportation as a speaker, whose private state is embedded under the speech of Undersecretary Harmann from sentence (5).

The **baseline** algorithm has a surprisingly good precision (see Tables 2 and3) but error analysis reveals some major problems. One is the assumption that the first-mentioned person would be the subject and, therefore, the speaker. The relatively flexible word order of German as well as passive-voice sentences cause this bet to fail in many instances such as (15), where the Secretary for Intra-German relations is the addressee of a request to be made but not himself a speaker.

(15)   a.   Der Bundesminister für gesamtdeutsche Fragen wird deshalb **gebeten**, im Ältestenrat darauf hinzuweisen, daß die Interpellation zumindest nicht vor der Beendigung der Außenministerkonferenz auf die Tagesordnung gesetzt wird.

         b.   *The Secretary for Intra-German relations will therefore be asked to suggest to the Council of Elders that the interpellation should at least not be put on the agenda before the end of the foreign ministers' conference.*

The baseline algorithm also loses recall by assigning only a single speaker, even though sentences can have multiple acts of communication and multiple speakers associated, as discussed above for sentence (6). Further, the simple string-matching method detects only 4 very frequent subjunctive mood forms, namely the third-person singular present forms *sei*, *werde*, and *habe* as well as third-person past form *würde*. It does not know if these forms really head the main clause or occur embedded, and it also misses

---

[4]See   http://www.semanticsoftware.info/reported-speech-tagger .

[5]German politics in the 1940s and 1950s was unfortunately entirely male-dominated and female speakers are sadly non-existent in our data set.

a lot of subjunctive forms of the same and other verbs. Finally, sentence-initial third person pronouns do not always signal speaker continuity. In (16), for instance, the referent of *er* is not a speaker but simply a topical referent that the speaker talks about.

(16) a. Er SEI in der Tat der einzige, der infolge seiner Personen- und Sachkenntnisse in der Lage sei, reformerische Maßnahmen durchzuführen.

  b. *He was the only one who could carry out reform measures based on his knowledge of the people and matters involved.*

Prec.	Loose	Strict
Baseline	77%	77%
Subject-based	81%	79%
Syntactic	86%$^{\ddagger\circ}$	69%†*
Recall	Loose	Strict
Baseline	44%	36%
Subject-based	65%‡	56%‡
Syntactic	87%‡*	79%‡*
F-score	Loose	Strict
Baseline	56%	49%
Subject-based	72%	65%
Syntactic	87%	74%

Table 2: Results on development set. $\dagger$ and $\ddagger$ indicate statistical significance compared to the baseline at $p < .05$ and $p < .01$ level, respectively; $*$ and $\circ$ indicate statistical significance compared to Subject-based method at $p < .01$ and $p < .05$ level, respectively (Fisher's Exact Test, two-tailed).

Prec.	Loose	Strict
Baseline	83%	83%
Subject-based	80%	79%
Syntactic	86%	72%†
Recall	Loose	Strict
Baseline	35%	35%
Subject-based	70%‡	70%‡
Syntactic	88%‡*	88%‡*
F-score	Loose	Strict
Baseline	49%	49%
Subject-based	75%	74%
Syntactic	87%	79%

Table 3: Results on test set. $\dagger$ and $\ddagger$ indicate statistical significance compared to the baseline at $p < .05$ and $p < .01$ level, respectively; $*$ indicates statistical significance compared to Subject-based method at $p < .01$ level (Fisher's Exact Test, two-tailed).

Our first algorithm following on the baseline is **subject-based** in that it addresses the problem that the first mention of a person in a sentence is not necessarily the subject by using the output of the Stanford parser (Klein & Manning 2003).[6] Our new algorithm works as follows:

---

1 If the current sentence s_i has a main clause subject go to step 2. Otherwise assign the person mentioned first in s_i as its speaker.

2 If the subject(s) occurring in s_i refer to persons from the biographical database, assign them as speakers. Otherwise, go to 3.

3 If s_i contains references to potential speakers, assign the first one as the subject. Otherwise, assign as speaker of s_i the speaker of s_{i-1}

Note that the subject-based algorithm ignores the subjunctive-mood indicators that the baseline uses as they turned out to deteriorate performance. This algorithm significantly outperforms the baseline on recall while having a comparable precision. The key to this boost is the fact that the current algorithm drastically decreases the number of sentences that are assigned "unknown" as speaker, which can only happen when s_i has no potential speaker and s_{i-1} does not either or there is no s_{i-1}. The baseline had 26687 cases of unknown speakers, whereas the subject-based algorithm has only 12914 instances. Note that the improvement would probably have been greater if the Stanford parser had been able to parse more of the long sentences in the protocols.

After another round of optimization, we arrive at a new algorithm which again makes use of subjunctive mood indicators. But instead of acting based on the presence of subjunctive forms anywhere in the sentence, we check for subjunctive mood only on main clause verbs. The final algorithm works as follows:

1 If current sentence s_i has a subjunctive mood main verb, assign speaker of s_{i-1}. Go on to 2

2 If s_i has a subject referring to potential speakers, add them to the set of speakers. If not, add the first-mentioned person in s_i to the set of speakers. Go on to 3.

3 If no speaker has been assigned so far, assign the speakers of s_{i-1}.

4 If the head verb is passive, assign the virtual speaker representing the cabinet as a whole.

The resulting algorithm (**syntactic**) is capable of cumulatively assigning multiple (sets of) speakers rather than just assigning one based on the first rule that fires. First, it checks for the presence of a subjunctive main verb as evidence for speaker continuity. Additionally, it can assign further speakers based on the evidence of the current sentence s_i, which allows us to correctly deal with embedded speakers. Specifically, the algorithm adds the subject referents of s_i if they are potential speakers according to the biographical database. If they are not, or if no parse is available, the algorithm adds the first-mentioned potential speaker in sentence s_i as a speaker, a fallback strategy similar to what Seki et al. (2009) use in their work on opinion holder identification. If no speaker is assigned at this stage, the algorithm simply carries over the speakers of the previous sentence s_{i-1}. Our final algorithm newly introduces

---

[6]http://nlp.stanford.edu/software/lex-parser.shtml

rule (4),, which is useful for picking up decisions by the cabinet, which are typically stated in the passive as illustrated in (17).

(17)   a.   Der vom Bundesminister der Justiz vorgetragene Entwurf wird antragsgemäß beschlossen.

   b.   *The bill presented by the Secretary of Justice is adopted as proposed.*

The new algorithm significantly increases recall compared to both the baseline and the subject-based method. On strict precision it performs somewhat worse than the other two methods, while on loose precision it scores higher than both the baseline and the subject-based method, though the difference is not statistically significant on the test set (Tables 2 and 3). With respect to the f-score the syntax-based method outperforms the other method noticeably.

Note that for information retrieval in a cultural heritage context, recall is typically more important than precision. For a historian searching for statements made by a particular politician it is important to be reasonably sure that all relevant information is found. If this cannot be guaranteed, i.e., if there are many false negatives, the search engine is of very limited use. On the other hand, erroneous information returned by the search engine, i.e., statements which were made by a different politician (false positives), is less harmful as they can be easily recognized and disregarded by the user.

6.   Conclusion

We presented a rule-based system for speaker attribution in cabinet protocols. We experimented with different heuristics and showed that it is possible to obtain a relatively high performance by exploiting linguistic cues.

While the performance of the system is fairly high, there are some remaining errors. For example, it is currently not possible to correctly identify embedded speech events, e.g., those evoked by a noun (e.g., "the report by ...") or preposition ("according to..."). We are currently working on addressing these shortcomings. One extension is to use a semantic role labeller (SRL) to find speakers based on their semantic role. The use of an SRL system also gives us information about speech events introduced by predicates other than verbs. Another extension is to use our rule-based system to label initial training data for a second stage supervised classifier, which can then exploit a larger set of linguistic cues to deal with the more difficult cases as well. Finally, we would also like to address topic identification as it is likely that not all speakers are equally likely to speak on any given topic.

Acknowledgements

7.   References

G. Atef, T. Barth, D. Quinten, A. Scheidel, and F. Shirokov. September 2009. SCOPE. Ein intelligentes webbasiertes Suchinterface für die Kabinettsprotokolle des Deutschen Bundestages. Unpublished project report.

S. Bergler. 1992. *The Evidential Analysis of Reported Speech.* Ph.D. thesis, Brandeis University, Boston, MA.

S. Bergler. 1995. Generative Lexicon Principles for Machine Translation: A Case for Meta-Lexical Structure. *Journal of Machine Translation*, 9(3):155–182.

S. Bethard, H. Yu, A. Thornton, V. Hatzivassiloglou, and D. Jurafsky. 2004. Automatic extraction of opinion propositions and their holders. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text*.

S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar. 2008. Building a sentiment summarizer for local service reviews. In *WWW Workshop on NLP in the Information Explosion Era (NLPIX)*.

Y. Choi, C. Cardie, E. Riloff, and S. Patwardhan. 2005. Identifying sources of opinions with conditional random fields and extraction patterns. In *Proceedings of EMNLP*, pages 355–362.

S.M. Kim and E. Hovy. 2006. Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, pages 1–8.

D. Klein and C.D. Manning. 2003. Fast Exact Inference with a Factored Model for Natural Language Parsing. In *Advances in Neural Information Processing Systems 15 (NIPS 2002)*, pages 3–10.

R. Krestel, S. Bergler, and R. Witte. 2008. Minding the source: Automatic tagging of reported speech in newspaper articles. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008)*.

Y. Seki, N. Kando, and M. Aono. 2009. Multilingual opinion holder identification using author and authority viewpoints. *Information Processing and Management*, 45:189–199.

J. Wiebe, T. Wilson, and C. Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39(2-3):165–210.

J. Wiebe. 1990. *Recognizing Subjective Sentences: A Computational Investigation of Narrative Text.* Ph.d., SUNY Buffalo Dept. of Computer Science, Buffalo.