Between Bags and Trees - Constructional Patterns in Text Used for Attitude Identification

Jussi Karlgren, Gunnar Eriksson, Magnus Sahlgren, and Oscar Täckström

Swedish Institute of Computer Science
Box 1263
SE-164 29 Kista
Sweden
{jussi,guer,mange,oscar}@sics.se
<http://www.sics.se>

Abstract. This paper describes experiments to use non-terminological information to find attitudinal expressions in written English text. The experiments are based on an analysis of text with respect to not only the vocabulary of content terms present in it (which most other approaches use as a basis for analysis) but also with respect to presence of structural features of the text represented by constructional features (typically disregarded by most other analyses). In our analysis, following a construction grammar framework, structural features are treated as occurrences, similarly to the treatment of vocabulary features. The constructional features in play are chosen to potentially signify opinion but are not specific to negative or positive expressions.

The framework is used to classify clauses, headlines, and sentences from three different shared collections of attitudinal data. We find that constructional features transfer well across different text collections and that the information couched in them integrates easily with a vocabulary based approach, yielding improvements in classification without complicating the application end of the processing framework.

1 Attitude Analysis is Mostly Based on Lexical Statistics

Attitude analysis, opinion mining, or sentiment analysis, a subtask of information refinement from texts, has gained interest in recent years, both for its application potential and for the promise of shedding new light on hitherto unformalised aspects of human language usage: the expression of attitude, opinion, or sentiment is a quintessentially human activity. It is not explicitly conventionalised to the degree that many other aspects of language usage are.

Most attempts to identify attitudinal expression in text have been based on lexical factors. Resources such as SentiWordNet, the Opinion Finder subjectivity lexicon, or the General Inquirer lexicon are utilised or developed by most research groups engaged in attitude analysis tasks [4, 18, 14]. But attitude is not a solely lexical matter. Expressions with identical or near-identical terms can be more or less attitudinal by virtue of their form (“He blew me off” vs. “He blew off”);

combinations of fairly attitudinally loaded terms may lack attitudinal power (“He has the best result, we cannot fail him” vs. “This is the best coffee, we cannot fail with it”); certain terms considered neutral in typical language use can have strong attitudinal loading in certain discourses or certain times (“Fifth Avenue”, “9/11”).

Our approach takes as its starting point the observation that lexical resources always are noisy, out of date, and most often suffer simultaneously from being both too specific and too general. Not only are lexical resources inherently somewhat unreliable or costly to maintain, but they do not cover all the possibilities of expression afforded by human linguistic behaviour.

We believe that attitudinal expression in text is an excellent test case for general purpose approaches for processing of linguistic data. We have previously tested resource-thrifty approaches for annotation of textual materials, arguing that general purpose linguistic analysis together with appropriate background materials for training a general language model provide a more general, more portable, and more robust methodology for extracting information from text [10].

This paper reports a series of experiments to investigate the general effectiveness of structural features as carriers of information in text, applied to the task of attitude analysis.

2 Constructions as Characteristic Features of Utterances

Our hypothesis is that investigating utterances for presence of content-bearing words may be useful for identifying attitudinal expressions, but that finding structural features carries over easier from one topical area to another, from one discourse to another.

It has previously been suggested that attitude in text is carried by dependencies among words, rather than by keywords, cue phrases, or high-frequency words [1]. We agree, but in contrast with previous work, we explicitly incorporate *constructions* in our knowledge representation, not as relations between terms but as features in their own right, following a construction grammar framework [9, 3].

Our claim is that the *pattern* of an utterance is a feature with the same ontological status as the terms that occur in the utterance: constructional features and lexical features both have conceptual meaning. Patterns are part of the signal, not incidental to it. This claim, operationalised for experimental purposes, gives us a convenient processing model. Where the step from bag-of-words analyses to complete parse trees is both computationally daunting and brittle in face of fluid and changing data, we can within a constructional framework find middle ground: we use observations of pattern occurrences as features similarly to how we use observations of word occurrences. An utterance will then not only be characterised as being a container for a number of words, but also a container for some observed patterns. Some previous approaches for using syntactic analysis in large-scale text analysis have used segments of parse trees rather than the entire tree; however, the distinction between lexical features indicating content

TENSE SHIFT *It **is** this, I **think**, that commentators **mean** mean when they **say** glibly
That the “world **changed**” after Sept 11.*

TIME ADVERBIAL *In Bishkek, they agreed to an informal meeting **later this year**,
most likely to be held in Russia.*

OBJECT CLAUSE *China could use the test as a political signal to show the US **that it
is a rising nuclear power** at this tense moment.*

VERB CHAIN *“Money **could be earned** by selling recycled first-run fuel and separated
products which retain over 50 per cent of unused uranium,” Interfax news agency
reported him as saying.*

Fig. 1. *K* examples.

and syntactic features indicating relations between lexical items are central in those analyses, while our approach does not separate those categories of features.

3 Combining Constructional and Lexical Features

The texts used in the present experiments are viewed as sequences of sentences: the sentence is taken as the basis of analysis, as a proxy for the utterance we view as the basis for attitudinal expression. All texts are preprocessed by a linguistic analysis toolkit, the Connexor Functional Dependency (FDG) parser, [17], to yield a lexical categorisation and morphological analysis of each word and a full dependency parse for each sentence. Our experimental features consist of three types: *content words* (*I*), *function words* (*F*) and *construction markers* (*K*) and are extracted from that analysis, making use of a mixture of low-level and abstract analysis results. Our features are general and not crucially bound to any specific analysis component, and no attitudinal lexical resources are employed to establish the constructions used in the further processes.

When observed in a sentence, the words from the *content word* class – nouns, adjectives, verbs (including verbal uses of participles), adverbs, abbreviations, numerals, interjections, and negation – form the feature class *I* for that sentence. Analogously, the *function words* – prepositions, determiners, conjunctions, pronouns – constitute the *F* feature class of the sentence.

Besides word occurrence based feature classes we introduce a further feature class intended to capture aspects of the constructions in employ in the sentence, *constructional features* (feature class *K*). Some markers, such as adverbial types and information about predicate and relative clauses are given directly by the FDG dependency or morphology analysis, others involved further processing steps to aggregate information from different levels of the analysis. The full list of *K* markers is given in Table 1.

The primary aim of the *K* features is to capture aspects of sentence composition, and therefore many features are concerned with clause types, and the way different types of clauses pattern in a sentence. Examples of such features,

Form	<i>K</i> tag	Freq	Form	<i>K</i> tag	Freq
Non-transitive clause	TRIN	2919	Adverbial of quantity	ADVLQUANT	85
Transitive clause	TRTR	2350	Undetermined		
Predicative clause	PREDCLS	1439	prepositional phrase	PPUNDET	22
Transitivity mix	TRMIX	1283	Negation	NEG	17
Tense shift	TNSSHIFT	733	Verb chain	VCHAIN	532
Object clause	OBJCLS	351	Verb particle	PARTV	7
Subordinating conjunction	SUBCNJ	200	Quantifier	QUANT	69
Relative clause	RELCLS	601	Adjective modifier	ADJMOD	82
Adverbial of location	ADVLSPAT	1367	Base form adjective	ABSADJ	3417
Adverbial of time	ADVLTIM	1110	Present tense	TNSPRES	2373
Sentence adverbial	ADVLSNT	973	Past tense	TNSPAST	2145
Adverbial of manner	ADVLMAN	608	Comparative adjective	KMPADJ	463
Adverbial of condition	ADVLCOND	547	Superlative adjective	SUPADJ	241
Clause-initial adverbial	ADVLCLSIN	387	Prepositional		
			post-modifier	PPPOMOD	572
			No K-traits	NOK	66

Table 1. *K*, with occurrence statistics from a corpus of approx. 90,000 words in 4,306 sentences of newspaper text.

involving a certain amount of further processing, are occurrence of transitive and intransitive clauses, and temporality pattern of clauses. The full list of *K* markers for sentence composition is: TRIN, TRTR, PREDCLS, TRMIX, TNSSHIFT, OBJCLS, RELCLS.

Another set of *K* features concerns the quality and type of adverbial, if present, in the sentence: ADVLSPAT, ADVLTIM, ADVLSNT, ADVLMAN, ADVLCOND, ADVLCLSIN, ADVLQUANT. Yet another type covers some morphological traits of sentences: the tense of verbs in clauses (obviously related to the tense pattern feature of the first subtype): TNSPRES and TNSPAST and the grade of occurring adjectives (ABSADJ, KMPADJ, and SUPADJ), which we expected to correlate well with our intended task to identify opinionated and attitudinal sentences. The remainder of the *K* feature set consists of markers of various word dependency relations, related to internal phrase structure, coordination, or clause polarity. These tags are: SUBCNJ, PPUNDET, NEG, PPPOMOD, VCHAIN, ADJMOD, QUANT, PARTV.

In the experiments below, all constructional features are treated as sentence features, exactly as the lexical features are treated, i.e., no coupling between the features and the words carrying them is made.

4 Test Data and Experimental Materials

For our experiments we use data from the NTCIR information retrieval evaluation challenge organised by NII, Tokyo, in its English section of the opinion identification task [12], the multi-perspective question answering (MPQA) test sentence

	"It is this, I think, that commentators mean when they say glibly that the 'world changed' after Sept 11."
<i>I</i>	be think commentator mean when say glibly world change sept 11
<i>F</i>	it this i that they that the after
<i>K</i>	ADVLTIM, ADVLMAN, OBJCLS, PREDCLS, TRIN, TRTR, TRMIX, TNSPRES, TNSPAST, TNSSHIFT
	"Mr Cohen, beginning an eight-day European tour including a Nato defence ministers' meeting in Brussels today and tomorrow, said he expected further international action soon, though not necessarily military intervention."
<i>I</i>	mr cohen begin eight-day european tour include nato defence minister meeting brussels today tomorrow say expect international action soon though not necessarily military intervention
<i>F</i>	an a in and he further
<i>K</i>	ADVLTIM, ADVLSPAT, OBJCLS, TRTR, TNSPAST

Fig. 2. Example attitude analyses of sentences. These sentence are taken from the NTCIR opinion analysis task data set. The first sentence is assessed by task judges to be a opinion carrier, the last non-opinion. The content word feature "say" is a strong marker for opinion but would yield the wrong categorisation in this case; our linear classifier correctly identified the first sentence as attitudinal and the last as non-attitudinal.

set with assessed attitudinal sentences [15], and the 2007 Semantic Evaluation Affective Task (SEMEVAL) test set of news headlines [16], all of which have assessments by human judges. We use a lenient scoring scheme, scoring a sentence as attitudinal if two out of three NTCIR judges have marked it attitudinal; for the SEMEVAL data if the intensity score is over 50 or under -50. All attitudinal sentences or headlines, irrespective of source, are assigned the class ATT and all other sentences assigned the class NOATT. Statistics for the collection are given in Table 2. Some sentences from the MPQA and NTCIR test sets, about ten in total, yielded no analyses in our system and are removed from the test set. These test sets are very different in character: the SEMEVAL set, which consists of news headlines rather than sentences from running text are different in structure; for each set, the assessors appear to have had different instructions. Our assumption is that this will approximate the variance that real-life tasks of this vein will encounter.

The data have been used by several research groups in various experiments. Our classifiers perform well enough to yield a tie with the reported best result from the shared opinion identification task of NTCIR as measured by F-score. (In the experiments described below we report precision and recall separately.)

5 Feature Strength Analysis

In order to gain some insight into which features show most utility for attitude identification we first performed some exploratory analyses on the NTCIR 6 and

	NTCIR 6	NTCIR 7	SEMEVAL	MPQA
Attitudinal	1392	1075	76	6021
Non-attitudinal	4416	3201	174	4982
Total	5808	4276	250	11003

Table 2. Test sentence statistics

7 test sets using NTCIR 6 as labeled training materials, testing on NTCIR 7 data, without using any other background material.

After training a SVM classifier for the ATT-NOATT distinction on all three (I , F , and K) feature types for the NTCIR data 2 245 features were utilised by the classifier. Most features were lexical, but we found that 17 K features were also used to discriminate between the classes. Table 3 list the 17 K markers in use with classifier-scored rank of importance, sub-grouping in the K set, and frequency of occurrence in the NTCIR-7 corpus. 85 of the remaining features were function words (F), and 2142 features were from the content word I set. The *rank* column in the table, given by the classifier, gives an indication of the relative importance of the features; the rank of feature set K in this given set is significantly higher than that of the I and F sets. (Mann Whitney U, $p > 0.95$).

We found that the relative scoring of the strongest features in the discrimination model scored certain of our manually chosen K features very highly compared to I and F features. Tense and transitivity measures, e.g., scored highly: “Tense shift”, the strongest single K feature is found in sentences where the verbs of the main clause and the subordinate clause have different tense forms. This occurs often in sentences of utterance or cognition: “Noam Chomsky said_{past} that what makes human language unique is_{present} recursive centre embedding”; “M.A.K. Halliday believed_{past} that grammar, viewed functionally, is_{present} natural”. The tense shift feature obviates the need of acquiring and maintaining lists of utterance, pronunciation, and cognition - categories which have obvious relation to attitudinal expression.

Another way to investigate the impact of different K traits is to study their occurrence in the sentences in the NTCIR-7 corpus. A matrix of K features and attitudinal status of sentences was constructed, and reduced to two dimensions by *correspondence analysis*, cf. [6]. This is a method similar to *principal components analysis*, but with the additional feature of placing the column and row variables on the same plane, and thus makes it possible to study the K features occurrence in sentences of different attitudinal type.

Figure 3 shows a plot of the result from the correspondence analysis, with non-opinionated outliers and less informative K labels ignored. The proximity of two labels is a measure of their co-occurrence, and we can notice that some K markers predominately show up in non-attitudinal sentences, e.g. verb chains and negation. On the opinionated side of the plane we find K traits as clause objects, tense shift patterns, adjectives in superlative grade, and predicative clauses. We

Rank	K tag	Type	Freq
75	TNSSHIFT	sentence composition	733
269	TNSPAST	morphology	2145
281	PARTV	phrase structure	7
290	TRMIX	sentence composition	1283
385	ADVLQUANT	adverbial	85
502	PREDCLS	sentence composition	1439
505	QUANT	phrase structure	69
680	TRIN	sentence composition	2919
686	NEG	polarity	17
746	ADVLTIM	adverbial	1110
780	TRTR	sentence composition	2350
813	PPPOMOD	phrase structure	572
969	PPUNDET	phrase structure	22
1055	KMPADJ	morphology	463
1105	VCHAIN	phrase structure	532
1673	TNSPRES	morphology	2373
2222	ADVLCOND	adverbial	547

Table 3. Ranking of K features according to decision function weights, among the 2245 predictive features in the NTCIR-7 corpus.

can also see that the opinionated and non-opinionated sentences are spread along the x-axis, the most important of the resulting two correspondence analysis dimensions, while the y-axis appears to involve the polarity of the sentence.

We have thus established that the K features carry signal value for predicting attitude in text. We now turn to investigating how much they might help in categorisation experiments. Rather than utilising the features directly from the test set, to avoid overtraining on the experimental sets we use a larger text collection to establish how similar the various features are in use.

6 Classification Experiment

Our main experiment is based on a background language representation built by analysis of a reasonable-sized general text collection. We use that model to establish similarities and differences between the sentences under analysis. Our general analysis procedure is to investigate how the utterance or sentence under consideration is related to language usage in the norm, either by deviation from the norm in some salient way, or by conforming with an identified model of usage.

In this experiment we use one year of newsprint from two Asian English-language news sources, the Korean Times and the Mainichi Daily, distributed as part of the NTCIR information retrieval evaluation challenge, including the opinion and attitude analysis task [7, 12]. We also use one year of the Glasgow Herald, distributed as part of the CLEF information retrieval evaluation challenge [2]. Collection sizes are as shown in Table 4.

Fig. 3. $K \times$ Attitude. Black labels denote opinionated sentences with positive (POS), negative (NEG), and neutral (NEU) polarity; NOATT denotes non-opinionated sentences.

We segment these corpora into sentences and process each sentence to extract the features given above – I , F , K . We use this to build a cooccurrence-based first-order word space [11] using the Random Indexing framework [8] (1000 dimensions, two non-zero elements per index vector).

Each *sentence* is represented by a random position in the word space model. Each observed *feature*, I , F , and K alike, is given an initially empty *context vector* in the word space model. Each feature observed in a sentence has its context vector incremented by vector addition with the representation vector for that sentence. Features which occur only once in the data were removed. Each feature will then in its context vector carry information about every sentence it has occurred in, and its context vector will eventually grow to become similar with other features it has cooccurred with. This is a standard word space model,

	Korean Times	Mainichi Daily	Glasgow Herald
Sentences	326 486	123 744	2 158 196
Characters	61M	25M	452M

Table 4. Background text materials

but here augmented with the K features. We use this word-and-feature space to be able to generalise from observed features in a sentence to other features and to establish similarities between sentences based on their feature values, even when there is little or no feature overlap.

Once the general language word-and-feature space model is built, each test sentence can be represented by the centroid of its feature set in the respective background space. Using these centroids, we used a support vector machine with a linear kernel to build a classifier for the ATT-NOATT distinction. The data were scaled to a range of approximately $[-1 \dots 1]$ and standard settings for the LIBLINEAR library [5] were used; since the class sizes were unbalanced, the penalty cost parameter for the classes was set in inverse proportion to their size.

We then ran five-fold crossvalidation of on each set to establish classification performance for the feature sets. This test was performed for each of the I , F , and K sets, and combinations thereof, yielding seven feature combinations for each set of test sentences and each background collection. The results are shown in Table 5 (precision), Table 6 (recall), and Table 7 (F_1).

We find that the combination of all three feature sets gives consistently high results as regards precision. We find that the including the K feature set gives the best results for recall for many of the combinations, and frequently using it alone gives the best results. We also find that the results are stable across the three background collections.

The test sets behave quite differently, however. News headlines (the SEMEVAL set) quite obviously give less purchase for classification. The low performance reported for the SEMEVAL set for the F feature set is unsurprising, since the language given in news headlines typically is quite terse, with structural cue words omitted for brevity, and rather different from the language in the background materials.

In comparison with other reported results these results are just slightly better than the best reported result for NTCIR-6 (best reported F_1 score from proceedings: 46.5) [13] and NTCIR-7 (best reported F_1 score from proceedings: 49.7) [12]; for the SEMEVAL task (best reported F_1 score from proceedings: 42.43) the best results from these experiments are comparable [16]. The MPQA corpus has been used in a large number of experiments and reported results vary, best F_1 scores being somewhat higher than the ones given here.

	NTCIR 6			NTCIR 7			MPQA			SMEVAL		
	KT	MD	GH	KT	MD	GH	KT	MD	GH	KT	MD	GH
<i>I</i>	37.7	36.0	37.5	37.4	35.4	37.4	71.9	68.9	70.7	25.4	24.6	33.9
<i>F</i>	35.0	35.4	39.0	35.4	35.2	39.9	70.3	69.8	72.3	20.0	25.0	31.8
<i>K</i>	33.6	33.7	37.1	35.3	35.2	37.2	66.4	66.2	71.1	25.4	26.4	26.7
<i>IF</i>	39.9	38.4	40.2	40.7	39.3	40.1	74.1	72.9	73.8	31.3	28.8	33.6
<i>IK</i>	38.5	37.2	40.2	39.7	38.5	40.9	71.8	69.7	73.7	25.0	23.9	30.5
<i>FK</i>	37.1	37.2	40.2	38.0	37.5	40.4	70.1	69.8	74.4	30.7	28.4	29.5
<i>IFK</i>	40.9	39.7	41.7	41.7	40.2	41.3	73.7	72.5	75.5	32.5	29.3	34.2

Table 5. Precision (%) for the NTCIR data set (five-fold crossvalidation).

	NTCIR 6			NTCIR 7			MPQA			SMEVAL		
	KT	MD	GH	KT	MD	GH	KT	MD	GH	KT	MD	GH
<i>I</i>	50.5	44.8	50.3	55.2	52.1	57.1	60.0	54.8	57.4	46.1	46.1	56.6
<i>F</i>	52.5	52.1	52.9	59.0	60.7	58.8	56.5	56.2	59.8	12.5	12.5	55.3
<i>K</i>	60.3	58.5	49.1	64.8	65.4	52.7	60.1	59.1	57.6	36.7	49.0	46.1
<i>IF</i>	54.3	52.1	53.5	59.2	58.0	58.0	61.5	58.1	61.9	55.3	52.6	54.0
<i>IK</i>	57.1	53.7	53.3	60.0	59.1	59.8	64.2	61.1	61.5	47.4	44.7	52.6
<i>FK</i>	55.8	54.6	54.0	59.9	59.3	58.9	58.9	58.3	62.6	51.3	52.6	51.3
<i>IFK</i>	59.4	56.6	55.3	62.3	61.6	59.1	64.2	61.4	63.9	51.3	53.9	54.0

Table 6. Recall (%) for the NTCIR data set (five-fold crossvalidation).

	NTCIR 6			NTCIR 7			MPQA			SMEVAL		
	KT	MD	GH	KT	MD	GH	KT	MD	GH	KT	MD	GH
<i>I</i>	43.2	39.9	46.1	44.6	42.2	45.2	65.4	61.0	63.4	32.7	32.1	42.4
<i>F</i>	42.0	42.2	44.9	44.2	44.6	47.5	62.6	62.3	65.4	15.4	16.7	40.4
<i>K</i>	43.2	42.8	42.3	45.7	45.7	43.6	63.1	62.4	63.7	30.0	34.3	33.8
<i>IF</i>	46.0	44.2	45.9	48.2	46.9	47.4	67.2	64.6	67.3	40.0	37.2	41.4
<i>IK</i>	46.0	44.0	45.9	47.8	46.6	48.6	67.8	65.1	67.0	32.7	31.2	38.6
<i>FK</i>	44.5	44.3	46.1	46.5	45.9	47.9	64.0	63.5	68.0	38.4	36.9	37.5
<i>IFK</i>	48.4	46.7	47.5	49.9	48.7	48.6	68.6	66.5	69.2	39.8	38.0	41.8

Table 7. F_1 (%) for the NTCIR data set (five-fold crossvalidation).

7 Conclusions

With increasingly sophisticated semantic relations being mined from data, processing must take a more sophisticated view of the linguistic signal than simple containers of topical words. Many approaches begin by assuming the structure of the linguistic data primarily to be relations between topical elements. In constructional approaches, the constructions, the combinational patterns, themselves are accorded presence in the signal – these experiments show that the patterns capture information, without the need for full analysis of dependencies between the word tokens. *Our level of analysis is between bags of words and parse trees in this respect.*

We find that representing constructions, even hand chosen constructions such as the ones given in this experiment, especially given unrelated general language background data, can provide a reliability which well matches or even surpasses that of word occurrence, the arguably primary carrier of information in the linguistic signal. The results are comparable to finely tuned experimental systems given by other research groups, trained for these specific experiments. *Constructions carry signal.*

Using constructions in parallel with word occurrence features not only has theoretical motivation based on the construction grammar theory, but also provides a convenient and familiar processing model and a straightforward extension for term based models. From a philological standpoint, a bottom-up approach to data analysis, examining the power of constructions as ontological items, would appear to be better motivated than basing information processing on descriptive language models, originally intended for description of human behaviour, for comparative studies of world languages, or for the scholarly instruction of foreign languages. *Constructions are easy to insert into a bag-of-words-based processing framework. Constructions do not promise more than they can deliver.*

Several of the constructions with greatest predictive effect for this example task were those that had the greatest scope: tense and transitivity patterns, predicative clauses. *Construction capture non-local information in sentences.*

The suggested K features might be one way of mimicking the contribution of constructions to the meaning of a sentence or utterance by combining these atomic markers of structure with the lexical items, without the need to represent the inter-relations between the two different sets. But the current implementation of the K traits idea has a number of drawbacks and limitations. The current set of K attributes was selected heuristically with task of attitude identification in mind. *Constructions should be learnt through data-driven methods, rather than selected heuristically.*

References

1. X. Bai, R. Padman, and E. Airoidi. On learning parsimonious models for extracting consumer opinions. In *Proceedings of HICSS-05, the 38th Annual Hawaii*

- International Conference on System Sciences*, Washington, DC, USA, 2005. IEEE Computer Society.
2. M. Braschler and C. Peters. Cross-language evaluation forum: Objectives, results, achievements. *Information Retrieval*, pages 7–31, 2004.
 3. W. Croft. Radical and typological arguments for radical construction grammar. In Östman and Fried [9].
 4. A. Esuli and F. Sebastiani. SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, 2006.
 5. R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008.
 6. M. J. Greenacre. *Theory and applications of correspondence analysis*. Academic Press, London, 1984.
 7. N. Kando. Overview of the Seventh NTCIR Workshop. In *Proceedings of the 7th NTCIR Workshop Meeting on Evaluation of Information Access Technologies*. NII, Tokyo, 2008.
 8. P. Kanerva, J. Kristofersson, and A. Holst. Random indexing of text samples for latent semantic analysis. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, page 1036. Erlbaum, 2000.
 9. J.-O. Östman and M. Fried, editors. *Construction Grammars: Cognitive grounding and theoretical extensions*. John Benjamins, Amsterdam, 2005.
 10. M. Sahlgrén, J. Karlgrén, and G. Eriksson. SICS: Valence annotation based on seeds in word space. In *Proceedings of Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, Czech Republic, 2007.
 11. H. Schütze. Word space. In S. Hanson, J. Cowan, and C. Giles, editors, *Advances in Neural Information Processing Systems 5*. Morgan Kaufmann Publishers, 1993.
 12. Y. Seki, D. K. Evans, L.-W. Ku, L. Sun, H.-H. Chen, and N. Kando. Overview of Multilingual Opinion Analysis Task at NTCIR-7. In *Proc. of the 7th NTCIR Workshop Meeting on Evaluation of Information Access Technologies*. NII, Tokyo, 2008.
 13. Y. Seki, D. K. Evans, L.-W. Ku, L. Sun, H.-H. Chen, N. Kando, and C.-Y. Lin. Overview of opinion analysis pilot task at ntcir-6. In N. Kando and D. K. Evans, editors, *Proceedings of the Sixth NTCIR Workshop Meeting on Evaluation of Information Access Technologies*. National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan, May 2007.
 14. P. Stone. Thematic text analysis: new agendas for analyzing text content. In C. Roberts, editor, *Text Analysis for the Social Sciences*, chapter 2. Lawrence Erlbaum Associates, 1997.
 15. V. Stoyanov, C. Cardie, D. Litman, and J. Wiebe, editors. *Evaluating an Opinion Annotation Scheme Using a New Multi-Perspective Question and Answer Corpus*. Stanford University, California, Mar. 2004. AAAI. AAAI Technical Report Series SS-04-07. Menlo Park: AAAI Press. ISSN 978-1-57735-219-x.
 16. C. Strapparava and R. Mihalcea. SemEval-2007 Task 14: Affective Text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 70–74, Prague, Czech Republic, June 2007. Association for Computational Linguistics.
 17. P. Tapanainen and T. Järvinen. A non-projective dependency parser. In *Proc. of the 5th Conference on Applied Natural Language Processing*, pages 64–71, 1997.
 18. T. Wilson, J. Wiebe, and P. Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of HLT-EMNLP*, 2005.