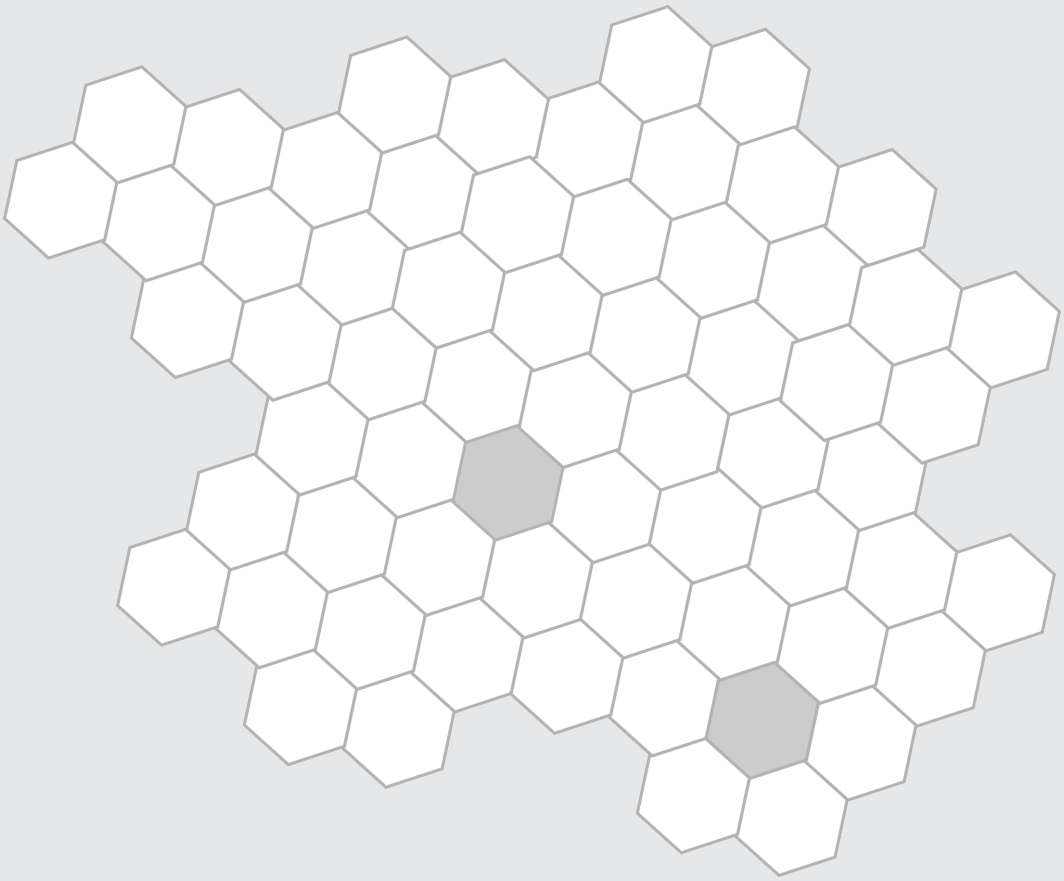
International Journal on

SOCIAL MEDIA

MMM: Monitoring, Measurement, and Mining



II, 2011, 1–2
ISSN 1804-5251

...OFFPRINT...

International Journal on SOCIAL MEDIA

MMM: Monitoring, Measurement, and Mining

II, 2011, 1–2

Editor-in-Chief: Jan Žižka

Publisher's website: www.konvoj.cz
E-mail: konvoj@konvoj.cz, SoNet.RC@gmail.com
ISSN 1804-5251

No part of this publication may be reproduced, stored or transmitted in any material form or by any means (including electronic, mechanical, photocopying, recording or otherwise) without the prior written permission of the publisher, except in accordance with the Czech legal provisions.

Linguistically-Motivated Subjectivity and Sentiment Annotation and Tagging of Modern Standard Arabic

Muhammad Abdul-Mageed

Department of Linguistics , and School of Library Information Science, Indiana University, Bloomington, 1320, E. Tenth Street Bloomington, IN, tel. U.S.A.+001-8128552018
e-mail: mabdulma@indiana.edu

Mona Diab

Center for Computational Learning Systems, Columbia University, NY, 475, Riverside Drive, Suite 850 Interchurch Center, MC 7177, tel. U.S.A.+001-2128701290
e-mail: mdiab@ccls.columbia.edu

Abstract

There has been recently a swelling interest in the area of *Subjectivity and Sentiment Analysis (SSA)*. However, only few attempts have been made to build SSA systems for *morphologically-rich languages (MRL)*. In addition, although there is a number of studies reporting annotation of various data sets for SSA, most of these studies are not strongly motivated by existing linguistic theory. In the current paper, we aim at partially bridging these two gaps in the literature. More specifically, we (1) provide a detailed description of a highly successful SSA system built for Modern Standard Arabic (MSA), a MRL, and (2) provide linguistically-motivated annotation guidelines for SSA.

Key words

subjectivity; sentiment, sentiment analysis; social meaning; Web mining; Arabic; morphologically-rich languages

Introduction

Subjectivity in natural language refers to aspects of language used to express opinions, feelings, evaluations, and speculations (Banfield, 1982; Wiebe, 1994) and, as such, it incorporates sentiment. The process of subjectivity classification refers to the task of classifying texts as either objective (e.g., *The NATO bombed Gaddafi troops*) or *subjective*. Subjective text is further classified with *sentiment* or *polarity*. For sentiment classification, the task refers to identifying whether a subjective text is *positive* (e.g., *The Egyptian and Tunisian revolutions are impressive!*), *negative* (e.g., The bloodbaths in Syria are horrifying!), *neutral* (e.g., *Saleh of Yemen may step down soon.*), or, sometimes, *mixed* (e.g., *I adore this camera, but it is prohibitively expensive*).

Two main issues arise in SSA. First, available approaches to labeling data for SSA vary considerably. Although our specific goal here is not to describe the various approaches, nor to provide a single standardized approach, we do seek to show how annotation studies within SSA can be inspired by existing linguistic theory. More specifically, by describing our efforts to label a specific corpus for SSA and

summarizing our linguistically-motivated guidelines for the task, we hope to trigger a stronger tie between existing linguistic theory and efforts to label data for social meaning tasks such as SSA.

Second, in spite of the flurry of research within the area of SSA, only few attempts have been made to build SSA systems for *morphologically-rich languages (MRL)* (i.e., languages in which significant information concerning syntactic units and relations are expressed at the word-level (Tsarfaty et al., 2010). We thus also aim at partially bridging this gap in the literature by reporting some of our recent work on Arabic, a very morphologically-complex language. We present work that investigates the role of morphology in SSA systems. We investigate Modern Standard Arabic (MSA), a morphologically-rich variety of Arabic, e.g., (Diab, 2007; Habash, Rambow and Roth, 2009). More specifically, we explore the task of sentence-level SSA on (MSA) texts from the news genre. We run experiments on three different pre-processing settings based on tokenized text from the Penn Arabic Treebank (PATB) (Maamouri et al., 2004) and employ both language-independent and Arabic-specific, morphology based features. Our work shows that explicitly using morphology-based features in our models improves the system's performance. We provide a detailed analysis of the performance of the system. We also measure the impact of using a wide coverage polarity lexicon and show that using a tailored resource results in significant improvement in classification performance.

The rest of the paper is organized as follows: In Section , we introduce the news genre. In Section , we overview the annotation process and the categories of each annotation task. In Section , we describe the (linguistics) insights driving our annotation. In Section , we provide examples from each SSA category. In Section , we describe our approach, including polarity lexicon, and automatic classification methodology. In Section , we present the results and evaluation. In Section , we provide an error analysis. Section is about related work and Section is the conclusion.

Subjectivity and Sentiment in the News

The bulk of SSA work has been performed on highly subjective, user-generated data such as blogs and product or movie reviews. In these data, authors tend to express their opinions quite explicitly Abdul-Mageed and Diab, 2011; Balahur and Steinberger, 2009). Consequently, data belonging to genres traditionally viewed as less subjective, like the news genre, have not received much attention. News, however, plays an instrumental role in modern societies (e.g., as an influencer of the social construction of reality Fowler, 1991; Chouliaraki and Fairclough, 1999; Wodak and Meyer, 2009).

In addition, news-making is increasingly becoming an interactive process in which lay Web users are getting more and more involved (Abdul-Mageed, 2008): News-makers reproduce some of the views of their readers (e.g., by quoting them) and they devote full stories to the interactions of web users on social media outlets. Although subjectivity in news articles has traditionally tended to be implicit, this growing trend to foster interactivity and more heavily report communication of internet users within the body of news articles is likely to make expression of subjectivity in news articles more explicit. Furthermore, the fact that news stories

	OBJ	S-POS	S-NEG	S-NEUT	Total
OBJ	1192	21	57	11	1281
S-POS	47	439	2	3	491
S-NEG	69	0	614	6	689
S-NEUT	115	2	9	268	394
Total	1423	462	682	288	2855

Table 1: Agreement for SSA sentences

have their own biases (e.g., hiding agents behind negative or positive events via use of passive voice, variation in lexical choice) has been pointed out (e.g., van Dijk, 1988) and hence the news genre is definitely not as objective as it may seem to some.

Data Annotation

Two graduate-level native speakers of Arabic annotated 2855 sentences from Part 1 V 3.0 of the Penn Arabic TreeBank (PATB) (Maamouri et al., 2004). The sentences make up the first 400 documents of that part of PATB amounting to a total of 54.5% of the PATB Part 1 data set. The task was to annotate MSA news articles at the sentence level. Each article has been processed such that coders are provided individual sentences to label. We prepared annotation guidelines for this SSA task focusing specifically on the newswire genre. We summarize the guidelines next, illustrating related and relevant literature.

Subjectivity and Sentiment Categories

For each sentence, each annotator assigned one of 4 possible labels: (1) Objective (OBJ), (2) Subjective-Positive (S-POS), (3) Subjective-Negative (S-NEG), and (4) Subjective-Neutral (S-NEUT). We followed Wiebe, Bruce and O'Hara (1999) in operationalizing the subjective vs. the objective categories. In other words, if the primary goal of a sentence is perceived to be the objective reporting of information, it was labeled OBJ. Otherwise, a sentence would be a candidate for one of the three subjective classes.[1] Table shows the contingency table for the two annotators' judgments. Overall agreement is 88.06%, with a Kappa (k) value of 0.38.

By way of illustration, a sentence such as "The Prime Minister announced that he will visit the city, saying that he will be glad to see the injured", has two authors (the story writer and the Prime Minister indirectly quoted). Accordingly to our guidelines, this sentence should be annotated S-POS tag since the part related to the person quoted (the Prime Minister) expresses a positive subjective sentiment,

[1] It is worth noting that even though some SSA researchers include subjective mixed categories, we only saw such categories attested in a negligible percent of the sentences which is expected since our granularity level is the sentence. If we are to consider larger units of annotation, we believe mixed categories will become more frequent. Thus we decided to tag the very few subjective mixed sentences as S-NEUT.

Domain	# of Cases
Politics	1186
Sports	530
Military & political violence	435
Disaster	228
Economy	208
Culture	78
Light news	72
Crime	62
This day in history	56
Total	**2855**

Table 2: Domains

"glad" which is a *private state* (i.e., a state that is not subject to direct verification) (Quirk et al., 1974).

Domain Annotation and Categories

The same two annotators also manually assigned each sentence a domain label. The domain labels are from the news genre and are adopted from (Abdul-Mageed, 2008). The set of domain labels is as follows: {*Light news, Military and political violence, Sport, Politics, Crime, Economy, Disaster, Arts and culture, This day in history*}. Table illustrates the number of sentences deemed for each domain. Domain annotation is an easier task than subjectivity annotation. Inter-annotator agreement for domain label assignment is at 97%. The two coders discussed differences and a total agreement was eventually reached. Coders disagreed most on cases belonging to the *Military and political violence* and *Politics* domains. For example, the following is a case where the two raters disagreed (and which was eventutally assigned a *Military and political violence* domain):

طلب رئيس الوزراء السابق في جزر فيدجي ماهندرا شودري الذي أطيح به

في ١٩ أيار مايو إثر حركة انقلابية، اليوم السبت باعادة حكومته إلى السلطة.

Transliteration: Tlb r}ys AlwzrA' AlsAbq fy jzr fydjy mAhndrA $wdry Al*y OTyH bh fy 19 OyAr mAyw Ivr Hrkp AnqlAbyp, Alywm Alsbt bIEAdp Hkwmth IlY AlslTp.
English: Former Prime Minister of Fiji Mahendra Chaudhry, who was ousted in May 19 after a revolutionary movement, asked on Saturday to return to office.

(Linguistic) Insights Driving Annotation

Good & Bad News vs. S-POS & S-NEG

In general, news can be either good or bad. For instance, whereas "Five persons were killed in a car accident" is bad news, "It is sunny and warm today in Chicago" is good news. Our coders were instructed not to consider *good* news *S-POS* nor *bad* news *S-NEG* if they think the sentences expressing them are objectively reporting information. Thus, bad news and good news can be OBJ as is the case in both examples. Indeed, this specific nuance makes news-focused SSA a difficult task.

Role of Epistemic Modality

Epistemic modality serves to reveal how confident writers are about the truth of the ideational material they convey (Palmer, 1986). Epistemic modality is classified into *hedges* and *boosters*. *Hedges* are devices like *perhaps* and *I guess* that speakers employ to reduce the degree of liability or responsibility they might face in expressing the ideational material. *Boosters*[2] are elements like *definitely, I assure that*, and *of course* that writers or speakers use to emphasize what they really believe. Both hedges and boosters can (1) turn a given unit of analysis from objective into subjective and (2) modify polarity (i.e., either strengthen or weaken it). Consider, for example, the sentences (1) "Gaddafi has murdered hundreds of people", (2) "Gaddafi may have murdered hundreds of people", and (3) "Unfortunately, Gaddafi has definitely murdered hundreds of people". While (1) is OBJ, since it lacks any subjectivity cues, (2) is S-NEUT because the proposition is not presented as a fact but rather is softened and hence offered as subject to counter-argument, (3) is a strong S-NEG (i.e., it is S-NEG as a result of the use of "unfortunately", and *strong* due to the use of the booster *definitely*). Our annotators were explicitly alerted to the ways epistemic modality markers interact with subjectivity.

Role of Perspective

Sentences can be written from different *perspectives* (Lin et al., 2006) or points of view. Consider the two sentences (1) "Israeli soldiers, our heroes, are keen on protecting settlers" and (2) "Palestinian freedom fighters are willing to attack these Israeli targets". While sentence (1) is written from an Israeli perspective, sentence (2) is written from a Palestinian perspective. The perspective from which a sentence is written interplays with how sentiment is assigned. Sentence (1) can usually be considered positive from an Israeli perspective, yet the act of protecting settlers is, more often than not, viewed as negative from a Palestinian perspective. Similarly, attacking Israeli targets may be positive from a Palestinian vantage point, but will perhaps be negative from an Israeli perspective. Coders were instructed to assign a tag based on their understanding of the type of sentiment, if any, the author of a sentence is trying to communicate. Thus, we have tagged the sentences from the perspective of their authors. As it is easy for a human to identify the perspective of an author (Lin et al., 2006), this measure facilitated the annotation task. Thus,

[2] Polanyi and Zaenen(2006) call these *intensifiers*.

knowing that the sentence (1) is written from an Israeli perspective, the annotator assigns it a S-POS tag.

Role of Illocutionary Speech Acts

Occurrences of language expressing (e.g. *apologies, congratulations, praise*, etc. are referred to as *illocutionary speech acts* (ISA) (Searle, 1976). We strongly believe that ISAs are relevant to the expression of sentiment in natural language. For example, the two categories *expressives* (e.g., congratulating, thanking, apologizing) and *commissives* (e.g., promising) of Searle's (1976) taxonomy of ISAs are specially relevant to SSA. In addition, Bach and Harnish (1979) define an ISA as a medium of communicating attitude and discuss ISAs like *banning, bidding, indicting, penalizing, assessing* and *convicting*. For example, the sentence "The army should never do that again" is a *banning* act and hence is S-NEG. Although our coders were not required to assign ISA tags to the sentences, we have brought the the concept of ISAs to their attention as we believe a good understanding of the concept does facilitate annotating data for SSA.

Role of Annotator's Background Knowledge

The type of sentiment expressed may vary based on the background knowledge of an annotator/reader (Balahur and Steinberger, 2009). For example, the sentence "Secularists will be defeated", may be positive to a reader who opposes secularism. However, if the primary intention of the author is judged to be communicating negative sentiment, annotators are supposed to assign a S-NEG tag. In general, annotators have been advised to avoid interpreting the subjectivity of text based on their own economic, social, religious, cultural, etc. background knowledge.

Examples of SSA categories from MSA news

We illustrate examples of each category in our annotation scheme. We also show and discuss examples for each category where the annotators differed in their annotations. Importantly, the two annotators discussed and adjudicated their differences.

Objective Sentences

Sentences where no opinion, sentiment, speculation, etc. is expressed are tagged as OBJ. Typically such sentences relay factual information, potentially expressed by an official source, like examples 1-3 below:

(١) ويبلغ عدد المشردين فى كنتية لوس انجلس نحو ٨٤ الف شخص.

Transliteration:[3] wyblg Edd Alm$rdyn fy kwntyp lws Onjlys nHw 84 Olf $xS.
English: The number of homeless in Los Angeles County is about 48 thousand.

[3] We use here Buckwalter transliteration www.qamus.org.

(٢) طهران ١٥ -٧ (أف ب) -وقع ١٦ انفجارا مساء اليوم السبت في وزارة الاستخبارات حيث استدعيت العديد من سيارات الاسعاف كما أكد شاهد عيا لوكالة فرانس برس.

Transliteration: ThrAn 15-7 (A f b) - wqE 16 AnfjArA msA' Alywm Alsbt fy wzArp AlAstxbArAt Hyv. AstdEyt AlEdyd mn syArAt AlIsEAf kmA Okd $Ahd EyAn lwkAlp frAns brs.
English: Tehran 15-7 (AFP) - An eye witness affirmed to AFP that 16 explosions occurred late Saturday at the Ministry of Intelligence where many ambulances were summoned.

(٣) أعلن السائق الأيرلندي أيدي أيرفاين (جاغوار) انسحابه من سباق جائزة النمسا الكبرى.

Transliteration: AEln AlsA}q AlIyrlndy Iydy IyrfAyn (jAgwAr) {nsHAbh mn sbAq jA}zp AlnmsA AlkbrY.
English: The Irish driver Eddie Irvine (Jaguar) announced his withdrawal from the Austrian Grand Prix.

Examples 1-3 show that objective sentences can have some implicitly negative words/phrases like "انتعشت الآمال{nsHAb" ("withdrawl"). In addition, although these 3 examples convey *bad* news, they are annotated with an OBJ tag since the sentences are judged as facts, although one annotator did initially tag example 1 as S-NEG before it was resolved later.

Subjective Positive Sentences

Sentences that were assigned a S-POS tag included ones with positive *private states* (Quirk et al., 1974; i.e., states that are not subject to verification). Examples 4 and 5 below are cases in point where the phrase "انسحاب AntE$t Al|mAl" ("hopes revived") and the word

اطمئنان

"TmnAn" ("relief") stand for unverifiable private states:

(٤) وانتعشت الامال بالافراج عن الرهائن في الساعات ال -- ٢٤ الاخيرة مع تدخل ليبيا.

Transliteration: wAntE$t Al*vert*mAl bAlIfrAj En AlrhA}n fy AlsAEAt Al 24 AlAxyrp mE tdxl lybyA.
English: Hopes for the release of hostages revived in the last 24 hours with the intervention of Libya.

(٥) و أبدى صلات حسن اطمئنانه إلى عودة النظام والاستقرار إلى بلاده.

Transliteration: wAbdY SlAt Hsn TmnAnh IlY Ewdp AlnZAm wAlstqrAr IlY blAdh.

English: Silaat Hasan expressed relief for the return of order and stability to his country.

The subtle nature of subjectivity as expressed in the news genre is reflected in some of the positive examples, especially in directly or indirectly quoted content when quoted people express their emotion or support their cause (via e.g., using modifiers). For instance, the use of the phrases من اجل نهضة الصومال"mn Ajl nhDp AlSwmAl" ("for the advancement of Somalia") in example 6 below turn what would have otherwise been OBJ sentences into S-POS sentences. Again, one annotator initially tagged example 8 as OBJ):

(٦) دعا الرئيس الصومالي مساء أمس السبت الدول المانحة وخصوصا أعضاء الجامعة العربية والاتحاد الأوروبي إلى تقديم مساعدات إلى بلاده "\من أجل نهضة الصومال "\.

Transliteration: dEA Alr}ys AlSwmAly msA' Ams Alsbt Aldwl AlmAnHp wxSwSA AEDA' AljAmEp AlErbyp wAl{tHAd AlAwrwby IlY tqdym msAEdAt IlY blAdh "mn Ajl nhDp AlSwmAl".
English: The Somali President, on Saturday evening, called on the donor countries, especially members of the Arab League and the European Union, to provide assistance to his country "for the advancement of Somalia".

Quoted content sometimes was in the form of *speech acts* (Searle, 1975). For example, (7) is an *expressive speech act* where the quoted person is thanking another party:

(٧) [وأضاف:] "\شكرا من أعماق قلبي لهذا الشرف الذي يمتد مدى الحياة. "\

Transliteration: [wADAf:] ""$krA mn AEmAq qlby lh'*A Al$rf Al*y ymtd mdY AlHyAp".
English: [He added:] Thank you from all my heart for this life-long honor.

Subjective Negative Sentences

Again, the more explicit negative content was found to be frequent in sentences with quoted content (as is illustrated in examples 8 and 9). (8) shows how the S-NEG S-POS sentiment can be very strong as is illustrated by the use of the noun phrase إصرار شيطاني "ISrAr $yTAny" ("diabolical insistence"):

(٨) ورد أحد محامي أندريوتي جيواكينو على قرار النيابة في باليرمو واصفا إياه بأنه "\إصرار شيطاني "\من قبل الاتهام.

Transliteration: wrd AHd mHAmy Andrywty jywAkynw sbAky ElY qrAr AlnyAbp fy bAlyrmw wASfA IyAh bAnh "ISrAr $yTAny" mn qbl AlAthAm.

English: One of lawyers of Andreotti Jjoaquino responded to the prosecutor's decision in Palermo, describing it as a "diabolical insistence" on the acusser's part.

Speech acts have also been used to express negative sentiment. For example, (14) is a direct quotation where a political figure denounces the acts of hearers. The speech act is intensified through the use of the adverb

(٩) وقال شارون من منصة الكنيست متوجها إلى نواب حزب العمل: "\لقد تخليتم حتى عن القسم الأكبر من المدينة القديمة "\.

"HtY" ("even"): حتى

Buckwalter: wqAl $Arwn mn mnSp Alknyst mtwjhA AlY nwAb Hzb AlEml "lqd txlytm HtY En Alqsm AlAkbr mn Almdynp Alqdymp."
English: Sharon, addressing Labour MPs from the Knesset, said: "You have even abandoned the biggest part of the old city".

Majority of the sentences pertaining to the *military and political violence* domain were OBJ, however, some of the sentences belonging to this specific domain were annotated S-NEG. News reporting is supposed to be objective, story authors sometimes used very negative modifiers, sometimes metaphorically as is indicated in (10). Example 10, however, was labeled OBJ by one of the annotators, and later agrrement was reached that it is more of an S-NEG case.

(١٠) وكان شهر تموز (يوليو) دمويا بشكل خاص مع سقوط نحو ٣٠٠ قتيل.

Transliteration: wkAn $hr tmwz ywlyw dmwyA b$kl xAS mE sqwT nHw 300 qtyl.
English: The month of July was especially bloody, with the killing of 300 people.

Subjective Neutral Sentences

Some of the S-NEUT cases were speculations about the future, as is illustrated by sentences 11 and 12:

(١١) ويتوقع أن يعود إلى الولايات المتحدة في ٢٥ تموز (يوليو).

Transliteration: wytwqE An yEwd IlY AlwlAyAt AlmtHdp fy 25 tmwz (ywlyw).
English: And he is expected to return to the United States on July 25.

(١٢) وكل الموءشرات تفيد أن هذا الوضع لن يتغير بعد الانتخابات.

Transliteration: wkl AlmW$rAt tfyd In h'*A AlwDE ln ytgyr bEd AlAntxAbAt.
English: All indications are that this situation will not change after the elections.

Hedges were also used to show cautious commitment to propositions, and hence turn OBJ sentences to S-NEUT ones. Sentences (13) and (14) are examples, with the occurrence of the hedge trigger word يبدو "ybdw" ("it seems") in (13) and على الأرجح "ElY AlArjH" ("it is most likely") in (14):

Word	POS	Surface form	Lemma	Stem	Gloss
AlwlAyAt	Noun	A+lwlAyAt	Al+wlAyp	Al+wlAy	the states
ltblgh	Verb	l+tblg+h	l+Oblg+h	l+blg+h	to inform him

Table 3: Examples of word lemmatization settings

(١٣) و يبدو أن التكتم الذي أحاط بزيارة بيريز إلى أندونيسيا كان يهدف إلى تفادي إثارة ردود فعل معادية في البلاد.

Transliteration: w ybdw An Altktm Al*y AHAT bzyArp byryz AlY AndwnysyA kAn yhdf AlY tfAdy AvArp rdwd fEl mEAdyp fy AlblAd.
English: It seems that the secrecy surrounding Peres's visit to Indonesia was aimed at avoiding negative reactions in the country.

(١٤) وعلى الأرجح أن قبطان الغواصة أعطى الأمر ب -إطفاء كل الآلات على متنها.

Transliteration: wElY AlArjH An qbTAn AlgwASp AETY AlAmr bATfA' kl AlAlAt ElY mtnhA.
English: Most likely the submarine's captain ordered turning off all the machines on board.

Approach

Polarity Lexicon

To the best of our knowledge, there is no publicly available polarity lexicon for Arabic. Accordingly, we manually created a lexicon of 3982 adjectives extracted from the first four parts of the PATB. Each adjective was labeled with one of the tags in the set {*positive, negative, neutral*}.

Automatic Classification Methodology

Settings: We run experiments on gold-tokenized text from PATB. We adopt the PATB+Al tokenization scheme, where proclitics and enclitics as well as the definite article "Al" are segmented out from the stem words. We experiment with three different pre-processing lemmatization configurations that specifically target the stem words: (1) *Surface*, where the stem words are left as is with no further processing of the morpho-tactics that result from the segmentation of clitics; (2) *Lemma*, where the stem words are reduced to their lemma citation forms, for instance in case of verbs it is the 3rd person masculine singular perfective form; and (3) *Stem*, which is the surface form minus inflectional morphemes. It is worth noting that the Stem configuration may result in non proper Arabic words (a la IR stemming). Table illustrates examples of the three configuration schemes, with each underlined.

Features: The features we employed are of two main types: Language-independent features and Morphological features.

Language-Independent Features:

This group of features has been employed in various SSA studies for English and other European Languages.

Domain: Following Wilson (2008), we apply a feature indicating the *domain* of the document to which a sentence belongs. As mentioned earlier, each sentence has a document domain label manually associated with it. The rationale behind applying this feature is that subjective language may be distributed differently across domains. For example, based on our development data, we notice that sentences belonging to the *Sports* domain tend to be more objective than those in, for example, the *Politics* domain.

UNIQUE: Following Wiebe et al. (2004), to meaningfully incorporate a feature that accounts for the frequency of words effect, we include a *unique* feature. Namely words that occur in our corpus with an absolute frequency ≤ 5 are replaced with the token "UNIQUE".

N-GRAM: We run experiments with N-grams ≤ 4 and all possible combinations of them.

ADJ: For subjectivity classification, we apply a binary *has_adjective* feature indicating whether or not any of the adjectives in our manually created polarity lexicon exists in a sentence. This is motivated by Bruce and Wiebe (1999) finding that adjectives are significantly and positively correlated with subjective sentences. For sentiment classification, we employ two features, *has_POS_adjective* and *has_NEG_adjective*, each of these binary features indicate whether a POS or NEG adjective occurs in a sentence.

MSA-Morphological Features: MSA exhibits a very rich morphological system that is templatic, agglutinative, and it is based on both derivational and inflectional features. We explicitly model morphological features of *person*, *state*, *gender*, *tense*, *aspect*, and *number*. We currently do not use part of speech information explicitly in our models. We assume undiacritized text in our models.

Method: Two-stage Classification Process

In the current study, we adopt a *two-stage* classification approach. In the first stage (*Subjectivity Classification*), we build a binary classifier to sort out OBJECTIVE from SUBJECTIVE cases. For the second stage (i.e., *Sentiment Classification*) we apply binary classification that distinguishes SUBJ-POS from SUBJ-NEG cases. We disregard the neutral class of SUBJ-NEUT for our current investigation. We use an Support Vector Machine classifier SVM$^{\text{light}}$ package (Joachims, 2008). We experiment with various kernels and parameter settings and find that linear kernels yield the best performance for our specific problem. We run experiments with *presence* vectors, i.e. for each sentence vector, the value of each dimension is binary either a 1 (regardless of how many times a feature occurs) or 0.

Experimental Conditions: We first run experiments using each of the three lemmatization settings *Surface, Lemma, Stem* using the various *N-GRAM* and *N-GRAM* combinations and then iteratively add other features exhaustively. Due to space lim-

	Surface form	Lemma	Stem
	F	F	F
	71.97	72.74	73.17
N-Gram	1g+2g+3g	1g+2g	1g+2g
Baseline	55.13	55.13	55.13

Table 4: Subjectivity Classification results on DEV data for the different lemmatization settings using N-GRAM features

itations, we present here the most interesting feature combinations namely: N-GRAM for all lemmatization settings, Stem+Morph, Lemma+ {DOMAIN, ADJ, UNIQUE}, Stem+Morph+{DOMAIN, ADJ, UNIQUE}.

With all the three settings, clitics that are split off words are kept as separate features in the sentence vectors.

Results and Evaluation

We divide our data into 80% for 5-fold cross-validation and 20% for test. For experiments on the test data, the 80% are used as training data. We have two settings, a development setting (DEV) and a test setting (TEST). In the development setting, we run the typical 5 fold cross validation where we train on 4 folds and test on the 5th and then average the results. In the test setting, we only ran with the best configurations yielded from the DEV conditions. In TEST mode, we still train with 4 folds but we test on the test data exclusively, averaging across the different training folds but crucially maintaining the same proportion of 4 folds for training. It is worth noting that the test data is larger than any given dev data (20% of the overall data set for test, vs. 16% for any DEV fold). We report results using *F*-measure (*F*).

Subjectivity

Table shows the overall results obtained in the DEV settings using N-GRAM features only. We experimented exhaustively with the different n-gram sizes and their combinations. As indicated in the table, different N-GRAM feature settings resulted in the best results for the different lemmatization configuration. In the Surface configuration, 1g+2g+3g N-GRAM feature combination yields the best results for this setting, while 1g+2g yields the best results for Lemma and Stem respectively. Stem yields the best classification results overall. All three results for Surface, Lemma, and Stem respectively outperform the Baseline which is a majority class baseline.

Furthermore, Table shows the impact of adding morphological features to the best yielding basic lemmatization setting Stem, as illustrated in the column Stem+Morph. Stem+Morph with N-GRAM feature 1g+2g+3g yields 73.48 *F* which is an increase over Stem without explicit morphological features that yields 73.17% *F*.

Adding language-independent features, illustrated in Table , the *ADJ* feature does not help in either the *Lemma* nor in the *Stem+Morph* settings, however, it helps

	Stem	Stem+Morph
	73.17	73.48
N-Gram	1g+2g	1g+2g+3

Table 5: Subjectivity Classification results on DEV data for Stem and Stem+Morph

	BASE	+ADJ	+DOMAIN	+UNIQUE
Lemma	72.74	72.12	72.64	73.05
Stem	73.17	73.22	72.78	**72.85**
Stem+Morph	**73.48**	**73.42**	**73.55**	72.53

Table 6: Subjectivity Classification results on DEV data for Lemma, Stem, and Stem+Morph +language-independent features

with the *Stem*. The *DOMAIN* feature improves the results only with the *Stem+Morph*. In addition, the *UNIQUE* feature modestly helps classification in the *Lemma* setting, but has a negative impact on both the *Stem* and the *Stem+Morph* settings. Table shows that although performance on TEST set drops with some settings on *Stem+Morph*, 6.25% improvement of *F* is acquired by applying the *ADJ* feature.

Sentiment

As Table shows, similar to the subjectivity results, the *Stem* setting performs better than the other two lemmatization settings, with 56.87% *F*, compared to 52.53% *F* for the *Surface* and 55.01% *F* for the *Lemma*, although this is still outperformed by the majority class baseline. Again, adding the morhology-based features helps improve the classification: The *Stem+Morph* is better than the *Stem* by about 1.00% *F*, as shown in Table . Table shows that whereas adding the *DOMAIN* feature helps in the *Lemma*, emphStem, and *Stem+Morph* settings, the *UNIQUE* feature only improves classification with the *Stem+Morph*. Adding the *ADJECTIVE* feature improves performance significantly: Improvements of 33.71% *F* for the *Lemma* setting, 34.06% *F* for the *Stem*, and 33.09% *F* for the *Stem+Morph* are possible. As Table shows, while performance on TEST data drops with application of the the *UNIQUE* feature, it slightly improves when the *DOMAIN* feature is added and significantly when the *ADJ* feature is used (the latter reaching 95.52% *F*).

Error Analysis

We performed an analysis of the errors made by the system in both the subjectivity and sentimment cases.

Stem+Morph	+ADJ	+DOMAIN	+UNIQUE
65.29	**71.54**	**63.54**	**65.17**

Table 7: Subjectivity Classification results on TEST data for Stem+Morph+language-independent features

	Surface form	Lemma	Stem
	F	F	F
	52.53	55.01	56.87
N-Gram	ig	ig	ig
Baseline	**58.65**	**58.65**	**58.65**

Table 8: Sentiment Classification results on DEV data for the different lemmatization settings using N-GRAM features

	Stem	Stem+Morph
	56.87	57.84
N-Gram	ig	ig

Table 9: Sentiment Classification results on DEV data for Stem and Stem+Morph

	BASE	+ADJ	+DOMAIN	+UNIQUE
Lemma	55.01	88.72	57.21	54.22
Stem	56.87	90.93	57.66	55.55
Stem+Morph	57.84	**90.93**	58.03	58.22

Table 10: Sentiment Classification results on DEV data for Lemma, Stem, and Stem+Morph +language-independent features

Stem+Morph	+ADJ	+DOMAIN	+UNIQUE
52.12	**95.52**	53.21	51.92

Table 11: Sentiment Classification results on TEST data for Stem+Morph+language-independent features

Error Analysis of Subjectivity Classification

The errors made by the system in the case of subjectivity analysis show that subjectivity is very context sensitive. Example 15 below, which was falsely classified by the system as subjective, illustrates that subjective words (e.g., "Alm'tqlyn" "the detainees", "AstslmwA" "they surrendered", "'frĝt" "it released", and "Astd'thm" "it issued summons to them") frequently occur in objective sentences and hence potentially confuse the classifier.

(١٥) وذكر مصدر في الشرطة أن أربعة عشر من المعتقلين السابقين ال ٣٥ في حركة حماس الذين أفرجت عنهم السلطة الفلسطينية استسلموا للشرطة الفلسطينية التي كانت استدعتهم.

English: A police source said that fourteen of the 35 former detainees in the Hamas movement, released by the Palestinian Authority, surrendered to Palestinian police, which had issued summons.
Buckwalter: w*kr mSdr fy Al$rTp In OrbEp E$r mn AlmEtqlyn AlsAbqyn Al 35 fy Hrkp HmAs Al*yn Ofrjt Enhm AlslTp AlflsTynyp AstslmwA ll$rTp AlflsTynyp Alty kAnt AstdEthm.

The system also failed to classify some of the sentences with time-constrained propositions (e.g., those where items like "ytwq'" "it is expected") occur and hence shift the class from factual to hypothetical/expected. Example 16 illustrates this specific case along with the fact that not only adjectives but polarized nouns (e.g., "mwAĝhAt" "confrontations" "qtlY" "killed people") as well can contribute to shifting a sentence class. Based on these specific observations, it might help to apply a feature indicating whether items capable of conditioning the time scope of propositions (e.g., "expected", "supposed") exist or not. Expanding the polarity lexicon beyond adjectives is also desirable.

(١٦) ويتوقع أن يلتقي الزعيمان الاسرائيلي إيهود باراك والفلسطيني ياسر عرفات غدا الاثنين في شرم الشيخ للتوصل إلى اتفاق حول وقف المواجهات التي أدت إلى سقوط ١٠٦ قتلى معظمهم من الفلسطينيين وأكثر من ثلاثة آلاف جريح.

English: The Israeli leader Ehud Barak and the Palestinian leader Yasser Arafat are expected to meet on Monday in Sharm el-Sheikh to reach agreement on ending the confrontations that led to the fall of 106 deaths, most of them Palestinians, and more than three thousand injured.
Buckwalter: wytwqE On yltqy AlzEymAn AlIsrA}yly Iyhwd bArAk wAlflsTyny yAsr ErfAt gdA AlAvnyn fy $rm Al$yx lltwSl IlY AtfAq Hwl wqf AlmwAjhAt Alty Odt IlY sqwT 106 qtlY mEZmhm mn AlflsTynyyn wOkvr mn vlAvp *vert*lAf jryH.

We also observed that most of the sentences belonging to the SPORT domain in our development set are objective. The system seemed to associate words from the sports domain more with the objective class and hence incidentally misclassifies

some sentences, as in example 17 below. Example 17 was misclassified as objective, even though it has polarized words. Observably, the example has more than one polarized items (e.g., "sĝlh hAfl" "track record", "'ahdr" "wasted", "_tmynp" "precious", "mrtf" "high"). It may thus prove useful to add a feature related to the number of polarized words used in a sentence.

(١٧) كثرت الترشيحات لمنتخبي الكويت وكوريا الجنوبية للمنافسة على لقب البطولة نظرا لسجلهما الحافل في البطولات السابقة، الأول لم يظهر بمستواه المعهود وأهدر نقطتين ثمينتين أمام إندونيسيا، والثاني كان مستواه مرتفعا جدا.

English: There have been many nominations for the national teams of Kuwait and South Korea to compete for the championship title because of of their track records in previous championships; the first did not keep up to its usual high level and wasted two valuable points against Indonesia, and the second showed a very high level.
Buckwalter: kvrt Altr$yHAt lmntxby Alkwyt wkwryA Aljnwbyp llmnAfsp ElY lqb AlbTwlp nZrA lsjlhmA AlHAfl fy AlbTwlAt AlsAbqp , AlOwl lm yZhr bmstwAh AlmEhwd wOhdr nqTtyn vmyntyn OmAm IndwnysyA, wAlvAny kAn mstwAh mrtfEA jdA.

Even though objective examples were observed to include numbers and reporting verbs more than subjective sentences, polarized content still co-existed with this objectivity cues (as is the case of example 18 below, which was misclassified as objective). A careful consideration of example 4 suggests that strongly polarized words (e.g., "mĝAzr" "massacres") can act as strong subjectivity cues. Our system does not currently have access to the strength of polarized words. Adding stength values to the polarity lexicon may improve the system's performance.

(١٨) وقد قتل أكثر من ٩٠ شخصا منذ بداية تشرين الأول أكتوبر في مجازر واعتداءات نسبت إلى الجماعات الاسلامية المسلحة المعادية لسياسة الوفاق الوطني التي يتبهجها الرئيّس عبد العزيز بوتفليقة وفق حصيلة استنادا إلى الصحف.

English: More than 90 people has been killed since the beginning of October in the massacres and attacks attributed to armed Islamist groups hostile to the national reconciliation policy pursued by President Abdelaziz Bouteflika, according to an outcome based on the newspapers.
Buckwalter: wqd qtl Okvr mn 90 $xSA mn* bdAyp t$ryn AlOwl Oktwbr fy mjAzr wAEtdA'At nsbt IlY AljmAEAt AlIslAmyp AlmslHp AlmEAdyp lsyAsp AlwfAq AlwTny Alty ynthjhA Alr}ys Ebd AlEzyz bwtflyqp wfq HSylp AstnAdA IlY Al-SHf.

Error Analysis of Sentiment Classification

The system misclassified some sentences whose adjective polarities are shift ed as a result of surrounding negation or polarity shifters (Polanyi and Zaenen, 2006).

Example 19 below has the adjective "dblwmAsyp" "diplomatic" preceded by the polarity shifter "qTE" "cutting", and was misclassified as positive. Applying a negation feature may thus improve the system's performance.

(١٩) وأضاف "\أول ترجمة لمثل هذا الدعم لا تكون إلا بقطع العلاقات الدبلوماسية وكل أشكال التبادل مع إسرائيل ومقاطعة إسرائيل على أوسع نطاق اقتصاديا "\.

English: He added, "[T]he first translation of such support can only happen by cutting off diplomatic relations and all forms of exchange with Israel ,and economically boycotting at on the scale".
Buckwalter: wODAf "Owl trjmp lmvl h'*A AldEm lA tkwn IlA bqTE AlElAqAt Aldblwmasyp wkl O$kAl AltbAdl mE IsrA}yl wmqATEp IsrA}yl ElY OwsE nTAq {qtSAdyA".

The system also did not cater for another category of polarity shifters (Polanyi and Zaenen, 2006), i.e., *epistemic modality* (with *hedges* like "rbmA" "perhaps" and *boosters* like "bAltOkyd" "certainly"). This resulted in errors like example 20 below where the hedging phrase "Al.hd mn" "reducing" softens the claim and hence alters the polarity.

(٢٠) وجاء في بيان نشر في ختام إجتماع لحكومة أن "\رئيس الوزراء أكد أمام الحكومة إن قمة شرم الشيخ تهدف إلى وقف العنف واقامة هيئة تعنى بالحد من مخاطر تجدد العنف ودراسة الأحداث التي وقعت منذ أسبوعين "\.

English: According to a statement released at the conclusion of a meeting of the government, "[T]he prime minister stressed to the government that the Sharm El-Sheikh summit aimed at halting the violence and establishing a body to reduce the risk of renewed violence and studying the events that took place two weeks ago."
Buckwalter: wjA' fy byAn n$r fy xtAm IjtmAE llHkwmp On {r}ys AlwzrA' Okd OmAm AlHkwmp In qmp $rm Al$yx thdf IlY wqf AlEnf wIqAmp hy}p tEnY bAlHd mn mxATr tjdd AlEnf wdrAsp AlOHdAv Alty wqEt mn* OsbwEyn".

The system also misclassified sentences with adjectives that were labeled neutral in the polarity lexicon, but which are polarized from the perspective (Lin et al., 2006) of the witer or person quoted in a news story. Example 21 below was missclassified as negative, although the adjective "Muslim" is positive from the perspective of the qouted person. Perspective identification may help assign polarities to certain adjectives.

(٢١) وأكد إن إيران "\ستساعد إقتصاديا شعب العراق المسلم "\.

English: He stressed that Iran would "help the Muslim people of Iraq economically."
Buckwalter: wOkd In IyrAn "stsAEd IqtSAdyA $Eb AlErAq Almslm".

The system also misclassified some sentences where multi-word polarized expressions (MWE) were used. Example 22 has an occurrence of a negative MWE (i.e., "Alal'b bAlnAr" "playing with fire"), but was misclassified as positive. This problem may be solvable by expanding the polarity lexicon with MWE.

(٢٢) وأخذ بيريز على الفلسطينيّن "\اللعب بالنار "\، وقال "\يجب أن يفهم الفلسطينيون أنهم يلعبون بالنار، ليس فقط مع إسرائيل بل مع العالم كله "\.

English: Peres balmed the Palestinians to for "playing with fire" and said, "Palestinians must understand that they are playing with fire, not only with Israel but with the whole world."
Buckwalter: wOx* byryz ElY AlflsTynyyn {AllEb bAlnAr}, wqAl {yjb On yfhm AlflsTynywn Onhm ylEbwn bAlnAr, lys fqT mE IsrA}yl bl mE AlEAlm klh"n.

Related Work

Regarding annotation of data for SSA, work on the news genre is most relevant to us. Wiebe, Wilson and Cardie (2005) describe a fine-grained news corpus manually labeled for SSA at the word and phrase levels. Their annotation scheme involves identifying the *source* and *target* of sentiment as well as other related properties (e.g., the *intensity* of expressed sentiment). Our work is less fine grained on the one hand, but we label our data for domain as well as subjectivity.

Balahur et al. (2009) report work on labeling quotations from the news involving one person mentioning another entity and maintain that quotations typically contain more sentiment expressions than other parts of news articles. Our work is different from that of Balahur et al. (2009) in that we label all sentences regardless whether they include quotations or not. Balahur et al. (2009) found that entities mentioned in quotations are not necessarily the target of the sentiment, and hence we believe that SSA systems built for news are better if they focus on all the sentences of articles rather than quotations alone (since the target of sentiment may be outside the scope of a quotation, but within that of the sentence to which a quotation belongs).

As for SSA systems, several sentence- and phrase-level classifiers have been built, (e.g., Wiebe, Bruce and O'Hara, 1999; Yi et al., 2003; Yu and Hatzivassiloglou, 2003; Kim and Hovy, 2004). Yi et al. (2003) present an NLP-based system that detects all references to a given subject, and determines sentiment in each of the references. Similar to Yi et al. (2003), Kim and Hovy (2004) present a sentence-level system that, given a topic detects sentiment towards it. Our approach differs from both Yi et al. (2003) and Kim and Hovy (2004) in that we do not detect sentiment toward specific topics. Also, we make use of *N*-gram features beyond unigrams and employ elaborate *N*-gram combinations.

Yu and Hatzivassiloglou (2003) build a document- and sentence-level subjectivity classification system using various *N*-gram-based features and a polarity lexicon. They report about 97% F-measure on documents and about 91% F-measure on sentences from the *Wall Street Journal* (WSJ) corpus. Some of our features are similar to those used by Yu and Hatzivassiloglou (2003), but we exploit additional

features. Wiebe, Bruce and O'Hara (1999) train a sentence-level probabilistic classifier on data from the WSJ to identify subjectivity in these sentences. They use POS features, lexical features, and a paragraph feature and obtain an average accuracy on subjectivity tagging of 72.17%. Again, our feature set is richer than Wiebe, Bruce and O'Hara (1999).

The only work on Arabic SSA we are aware of is that of Abbasi, Chen and Salem (2008). They use an entropy weighted genetic algorithm (EWGA) for both English and Arabic Web forums at the document level. They exploit both syntactic and stylistic features. Abbasi et al. use a root extraction algorithm and do not use morphological features. They report 93.6% accuracy. Their system is not directly comparable to ours due to the difference in data sets and tagging granularity.

Conclusion

In this paper, we present a novel annotation layer of SSA to an already labeled MSA data set, the PATB Part 1 ver. 3.0. To the best of our knowledge, this layer of annotation is the first of its kind on MSA data of the newswire genre. We will make that collection available to the community at large. We motivate SSA for news and summarize our linguistics-motivated guidelines for data annotation and provide examples from our data set. We also build a sentence-level SSA system for MSA contrasting language independent only features vs. combining language independent and language-specific feature sets, namely morphological features specific to Arabic. We also investigate the level of stemming required for the task. We show that the *Stem* lemmatization setting outperforms both *Surface* and *Lemma* settings for the SSA task. We illustrate empirically that adding language specific features for MRL yields improved performance. Similar to previous studies of SSA for other languages, we show that exploiting a polarity lexicon has the largest impact on performance. We also identify several areas where the system makes errors, with a view to future improvement.

References

Abbasi, A., Chen, H., Salem, A. (2008): Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums. *ACM Trans. Inf. Syst.*, Vol. 26, pp. 1–34.

Abdul-Mageed, M. (2008): Online news sites and Journalism 2.0: Reader comments on Al Jazeera Arabic. *Triple C: Cognition, Communication, Cooperation*, Vol. 6, No. 2, pp. 59–76.

Abdul-Mageed, M., Diab, M. (2011): Subjectivity and Sentiment Annotation of Modern Standard Arabic Newswire. In: *LAW V : Fifth Linguitic Annotation Workshop : Proceedings of the Workshop*. Stroudsburg (Philadelphia, USA) : Association for Computational Linguistics, pp. 110–118. Available at: `http://www.aclweb.org/anthology/W11-0413`.

Bach, K., Harnish, R. M. (1979): *Linguistic communication and speech acts*. Cambridge (MA) : MIT Press, 327 pp.

Balahur, A., Steinberger, R. (2009): Rethinking Sentiment Analysis in the News: from Theory to Practice and Back. In: *WOMSA '09 : Proceedings of the 1st Workshop on Opinion Mining and Sentiment Analysis*. Sevilla : University of Sevilla, pp. n/a. Available at: `http://130.203.133.150/viewdoc/versions?doi=10.1.1.157.3642`.

Balahur, A., Steinberger, R., van der Goot, E., Pouliquen, B., Kabadjov, M. (2009): Opinion Mining on Newspaper Quotations. In: *2009 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. IEEE Computer Society, pp. 523–526. (ISBN 978-0-7695-3801-3.)

Banfield, A. (1982): *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Boston : Routledge & Kegan Paul, 340 pp.

Bruce, R., Wiebe, J. (1999): Recognizing subjectivity. A case study of manual tagging. *Natural Language Engineering*, Vol. 5, No. 2, pp. 187–205.

Diab, M., Hacioglu, K., Jurafsky, D. (2007): Automatic processing of Modern Standard Arabic text. In: *Arabic Computational Morphology.* Heidelberg : Springer, pp. 159–179. (ISBN 978-1-4020-6046-5.)

Fowler, R. (1991): *Language in the News: Discourse and Ideology in the Press.* London : Routledge, 254 pp. ISBN 0-415-01418-2.

Habash, N., Rambow, O., Roth, R. (2009): MADA+TOKAN: A Toolkit for Arabic Tokenization, Diacritization, Morphological Disambiguation, POS Tagging, Stemming and Lemmatization. In: Choukri, Kh., Maegaard, B. *Proceedings of the Second International Conference on Arabic Language Resources and Tools* [MEDAR'09]. [s. l.] : MEDAR Consortium, pp. 102–109. (ISBN 2-9517408-5-9.)

Chouliaraki, L., Fairclough, N. (1999): *Discourse in late modernity : Rethinking critical discourse analysis.* Edinburgh : Edinburgh University Press, 168 pp. ISBN 0-7486-1082-0.

Joachims, T. (2008): SVM^{light} : Support Vector Machine v. 6.02 [software].

Bach, K., Harnish, R. M. (1979): *Linguistic communication and speech acts.* Cambridge (MA) : MIT Press, 327 pp.

Kim, S., Hovy, E. (2004): Determining the sentiment of opinions. In: *Coling 2004 : 20th International Conference on Computational Linguistics : Proceedings of the Conference.* Stroudsburg (Philadelphia, USA) : Association for Computational Linguistics, pp. 1367–1373.

Lin, W. H., Wilson, T., Wiebe, J., Hauptmann, A. (2006): Which side are you on?: identifying perspectives at the document and sentence levels. In: *Proceedings of the Tenth Conference on Computational Natural Language Learning.* Stroudsburg (Philadelphia, USA) : Association for Computational Linguistics, pp. 109–116.

Maamouri, M., Bies, A., Buckwalter, T., Mekki, W. (2004): The penn arabic treebank: Building a large-scale annotated arabic corpus. In: *NEMLAR Conference on Arabic Language Resources and Tools.* , pp. 102–109.

Palmer, F. (1986): *Mood and Modality. 1986.* Cambridge : Cambridge University Press, x+243 pp.

Polanyi, L., Zaenen, A. (2006): Contextual valence shifters. In: *Computing attitude and affect in text: Theory and applications.* Heidelberg : Springer, pp. 1–10.

Quirk, R., Greenbaum, S., Close, R. A., Quirk, R. (1974): *A university grammar of English.* London : Longman, 484 pp.

Searle, J. R. (1975): A taxonomy of speech acts. In: Gunderson, K. *Language, mind, and knowledge.* Minneapolis : University of Minnesota Press, pp. 344–369.

Tsarfaty, R., Seddah, D., Goldberg, Y., Kuebler, S., Versley, Y., Candito, M., Foster, J., Rehbein, I., Tounsi, L. (2010): Statistical Parsing of Morphologically Rich Languages (SPMRL) What, How and Whither. In: *Proceedings of the NAACL HLT 2010 First Workshop on Statistical Parsing of Morphologically-Rich Languages.* Stroudsburg (Philadelphia, USA) : Association for Computational Linguistics, pp. 1–12.

Van Dijk, T. A. (1988): *News as discourse.* Mahwah (NJ) : Lawrence Erlbaum Associates, 200 pp.

Wiebe, J. (1994): Tracking point of view in narrative. *Computional Linguistics*, Vol. 20 No. 2, pp. 233–287.

Wiebe, J., Bruce, R., O'Hara, T. (1999): Development and use of a gold standard data set for subjectivity classifications. In: *37th Annual Meeting of the Association for Computational Linguistics : Proceedings of the Conference.* Stroudsburg (Philadelphia, USA) : Association for Computational Linguistics, pp. 246–253. (ISBN 1-55860-609-2.)

Wiebe, J., Wilson, T., Bruce, R., Bell, M., Martin, M. (2004): Learning subjective language. *Computational Linguistics*, Vol. 30, No. 3, pp. 277–308. (ISSN 0891-2017)

Wiebe, J., Wilson, T., Cardie, C. (2005): Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, Vol. 39, No. 2, pp. 165–210. (ISSN 1574-020X)

Wilson, T., Wiebe, J., Hoffmann, P. (2009): Recognizing Contextual Polarity: an exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, Vol. 35, No. 3, pp. 399–433. (ISSN 0891-2017)

Wodak, R., Meyer, M. (2009): *Methods of critical discourse analysis.* London : Sage, pp. 1–33. Chapter 1: Critical discourse analysis: History, agenda, theory and methodology.

Yi, J., Nasukawa, T., Bunescu, R., Niblack, W. (2003): Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In: *Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM 2003).* IEEE Computer Society, pp. 427–434. (ISBN 0-7695-1978-4.)

Yu, H., Hatzivassiloglou, V. (2003): Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In: *Proceedings of the 2003 conference on Empirical methods in natural language processing. Vol. 10 (EMNLP'03).* Stroudsburg (Philadelphia, USA) : Association for Computational Linguistics, pp. 129–136.