Sentiment Composition

Karo Moilanen and Stephen Pulman
Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford, OX1 3QD, England
{ *Karo.Moilanen* | *Stephen.Pulman* }@*comlab.ox.ac.uk*

Abstract

Sentiment classification of grammatical constituents can be explained in a quasi-compositional way. The classification of a complex constituent is derived via the classification of its component constituents and operations on these that resemble the usual methods of compositional semantic analysis. This claim is illustrated with a description of sentiment propagation, polarity reversal, and polarity conflict resolution within various linguistic constituent types at various grammatical levels. We propose a theoretical composition model, evaluate a lexical dependency parsing post-process implementation, and estimate its impact on general NLP pipelines.

Keywords

sentence-level sentiment, clause-level sentiment, entity-level sentiment, valence shifters, polarity shifters, lexical semantics

1 Introduction

Using lists of positive and negative keywords can give the beginnings of a sentiment classification system. However, classifying sentiment on the basis of individual words can give misleading results because atomic sentiment carriers can be modified (weakened, strengthened, or reversed) based on lexical, discoursal, or paralinguistic contextual operators ([7]). Past attempts to deal with this phenomenon include writing heuristic rules to look out for negatives and other 'changing' words ([6]), combining the scores of individual positive and negative word frequencies ([11], [5]), and training a classifier on a set of contextual features ([10]). While statistical sentiment classifiers work well with a sufficiently large input (e.g. a 750-word movie review), smaller subsentential text units such as individual clauses or noun phrases pose a challenge. It is such low-level units that are needed for accurate entity-level sentiment analysis to assign (local) polarities to individual mentions of people, for example.

In this paper we argue that, as far as low-level (sub)sentential sentiment classification is concerned, there may be much to be gained from taking account of more linguistic structure than is usually the case. In particular we argue that it is possible to calculate in a systematic way the polarity values of larger syntactic constituents as some function of the polarities of their subconstituents, in a way almost exactly analogous to the 'principle of compositionality' familiar from the formal semantics literature ([2]). For if the meaning of a sentence is a function of the meanings of its parts then the global polarity of a sentence is a function of the polarities of its parts. For example, production rules such as [$VP_\alpha \rightarrow V_\alpha + NP$] and [$S_\beta \rightarrow NP + VP_\beta$] operating on a structure like *"America invaded Iraq"* would treat the verb *"invade"* as a function from the NP meaning to the VP meaning (i.e. as combining semantically with its direct object to form a VP). The VP meaning is correspondingly a function from the NP meaning to the S meaning (i.e. as combining with a subject to form a sentence). Analogously, a 'DECREASE' verb like *"reduce"* (cf. [1]) should then be analysed as having a compositional sentiment property such that it reverses the polarity (α) of its object NP in forming the VP, hence [$VP_\beta^{(\neg\alpha)} \rightarrow V_{\beta[DECREASE]} + NP^{(\alpha)}$]. Thus the positive polarity in *"reduce the risk"* even though *"risk"* is negative in itself (cf. the negative polarity in *"reduce productivity"*). In fact, this semi-compositionality also holds at other linguistic levels: certainly amongst morphemes, and arguably also at suprasentential levels. However, this paper discusses only sentential sentiment composition. Grounded on the descriptive grammatical framework by ([4]), we propose a theoretical framework within which the sentiment of such structures can be calculated.

2 Composition Model

The proposed sentiment composition model combines two input (IN) constituents at a time and calculates a global polarity for the resultant composite output (OUT) constituent (cf. parent node dominance in the *modifies_polarity* and *modified_by_polarity* structural features in ([10])). The two IN constituents can be of any syntactic type or size. The model assumes dominance of non-neutral (positive (+), negative (-), mixed (M)) sentiment polarity over neutral (N) polarity. The term **sentiment propagation** is used here to denote compositions in which the polarity of a neutral constituent is overridden by that of a non-neutral constituent ({(+)(N)} → (+); {(-)(N)} → (-)). We use the term **polarity reversal** to denote compositions in which a non-neutral polarity value is changed to another non-neutral polarity value ((+) → (-); (-) → (+)) (cf. [7]), and the term **polarity conflict** to denote compositions containing conflicting non-neutral polarities ({(+)(-)} → (M)). Polarity **conflict resolution** refers to disambiguating compositions involving a polarity conflict ((M) → (+); (M) → (-)).

Polarity conflict resolution is achieved by ranking the IN constituents on the basis of relative weights assigned to them dictating which constituent is more important with respect to sentiment. The stronger of the IN constituents is here denoted as <u>SPR</u> (superordinate) whereas the label SUB (subordinate) refers to the dominated constituent (i.e. <u>SPR</u> $\gg$ SUB). Except for (N)[=] <u>SPR</u> constituents, it is therefore the <u>SPR</u> constituent and the compositional processes executed by it that determine the polarity (α) of the OUT constituent (i.e. $\text{OUT}^{\alpha_{ij}} \rightarrow \underline{\text{SPR}}^{\alpha_i} + \text{SUB}^{\alpha_j}$). The weights are not properties of individual IN constituents per se but are latent in specific syntactic constructions such as `[Mod:Adj Head:N]` (e.g. adjectival premodification of head nouns) or `[Head:V Comp:NP]` (e.g. direct object complements of verbs).

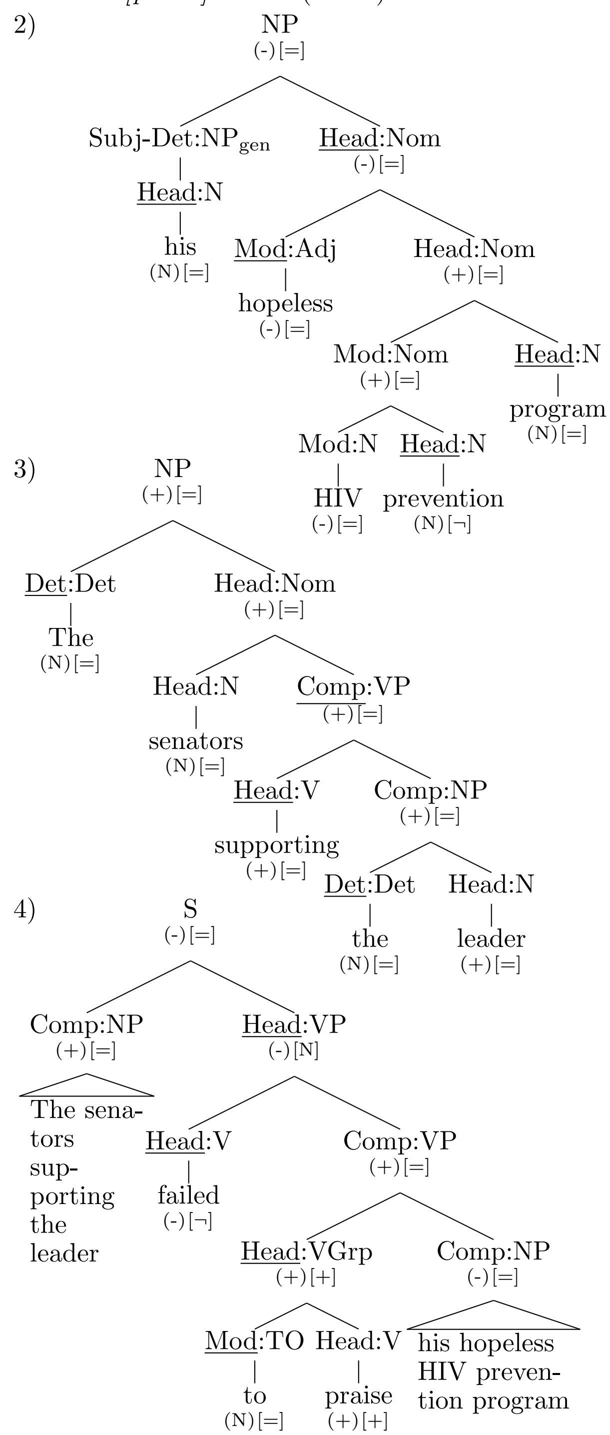
We tag each entry in the sentiment lexica (across all word classes) and each constituent with one of the following tags: **default** ([=]), **positive** ([+]), **negative** ([-]), and **reverse** ([¬]). These tags allow us to specify at any structural level and composition stage what any given <u>SPR</u> constituent does *locally* to the polarity of an accompanying SUB constituent without fixed-order windows of n tokens (cf. ([7]), modification features in ([10]), change phrases in ([6])). A [=] <u>SPR</u> constituent combines with a SUB constituent in the default fashion. The majority of constituents are [=]. A [¬] <u>SPR</u> constituent reverses the polarity of the SUB constituent and assigns that polarity to the OUT constituent (cf. general polarity shifters in ([10])). As <u>SPR</u> constituents, some carriers such as *“[contaminate]$^{(-)}$”* or *“[soothe]$^{(+)}$”* exhibit such strong sentiment that they can determine the OUT polarity irrespective of the SUB polarity - consider the static negativity in *“[contaminated that damn disk]$^{(-)}$”*, *“[contaminated the environment]$^{(-)}$”*, and *“[contaminated our precious water]$^{(-)}$”* (vice versa for some positive carriers). Hence the [-] and [+] constants which can furthermore be used as polarity heuristics for carriers occurring prototypically with a specific polarity (e.g. *“[deficiency (of sth positive)]$^{(-)}$”*) (cf. presuppositional items in ([7]), negative and positive polarity shifters in ([10])).

Notice that the <u>SPR</u> constituent operates on the SUB constituent irrespective of the polarity of the latter as a [¬] <u>SPR</u> constituent such as the determiner *“[less]$^{(N)[\neg]}$”* reverses both (+) and (-) SUB constituents (e.g. *“[less tidy]$^{(-)}$”*, *“[less ugly]$^{(+)}$”*), for example. However, cases in which <u>SPR</u> operations are required only in conjunction with a specific SUB constituent polarity do exist. The reversal potential in the degree modifier *“[too]$^{(N)[\neg]}$”*, for instance, seems to operate only alongside (+) SUB constituents (i.e. *“[too colourful]$^{(-)}$”* vs. *“??[too sad]$^{(+)}$”*). The adjective *“[effective]$^{(+)[=]}$”* operates similarly only with (+) or (N) SUB constituents (i.e. *“[effective remedies/diagrams]$^{(+)}$”* vs. *“[effective torture]$^{(-)}$”*). It is thus proposed that (?:+) and (?:-) be used as further **filters** to block specific SPR polarities as required by individual carriers.

To illustrate how the composition model operates, consider the sample sentence in Ex. 1:

1) *The senators* ***supporting***$^{(+)}$ *the* ***leader***$^{(+)}$ ***failed***$^{(-)}$ *to* ***praise***$^{(+)}$ *his* ***hopeless***$^{(-)}$ ***HIV***$^{(-)}$ *prevention program.*

Raw frequency counts, yielding three (+) and three (-) carriers, would fail to predict the global negative polarity in the sentence. We represent the sentence as follows, starting with the direct object NP of the predicator *“[praise]$^{(+)[+]}$”* (Ex. 2):



Through polarity reversal, the internal sentiment in *“[HIV prevention]$^{(+)[=]}$”* is first arrived at due to the [¬] status of the <u>SPR</u> head noun *“[prevention]$^{(N)[\neg]}$”* which reverses the (-) premodifying noun *“[HIV]$^{(-)[=]}$”*. The (N) head noun *“[program]$^{(N)[=]}$”* is then overridden by the (+) premodifying nominal *“[HIV prevention]$^{(+)[=]}$”*. When the resultant nominal is combined with the premodifying attributive <u>SPR</u> input *“[hopeless]$^{(-)[=]}$”*, the ensuing polarity conflict can be resolved through the dominance of the premodifier in this syntactic situation. The final combination with the SUB subject determiner *"[his]$^{(N)[=]}$"* is a case of propagation as the resultant NP reflects the polarity of the head nominal. Sentiment propagation can be seen throughout the subject NP (Ex. 3) as the (+) head noun *"[leader]$^{(+)[=]}$"*, combined with a (N) SPR determiner, results in a (+) NP (*"[the leader]$^{(+)[=]}$"*). When that NP is combined with a (+) SPR head participial, a (+) SPR VP is generated (*"[supporting the leader]$^{(+)[=]}$"*) which in turn overrides the (N) head noun *"[senators]$^{(N)[=]}$"*. The final (N) SPR determiner does not change the polarity any further.

The NPs thus resolved can then be combined with the two predicators to form a sentence (Ex. 4). The direct object NP *"[his hopeless HIV prevention program]$^{(-)[=]}$"* is reversed when it is combined with an SPR verb group outputting constant positivity (*"[to praise]$^{(+)[+]}$"*). When the resultant (+) VP is used as the complement of a [¬] SPR head verb polarity reversal occurs once again yielding a (-) VP (*"[failed to praise his hopeless HIV prevention program]$^{(-)[=]}$"*). Lastly, the (+) subject NP combines with the (-) predicate, and the polarity conflict is resolved due to the predicate being the SPR constituent. Hence the global negative sentiment for the present sample sentence can be calculated from its subconstituents.

3 Grammatical Constructions

Within a syntactic phrase, the polarity of the phrasal head can be changed by its pre- and post-modifying dependents. In general, pre-head dependents dominate their heads. **Determiners** (e.g. *"[no crime]$^{(-)}$"*) and **DPs** (e.g. *"[too much wealth]$^{(-)}$"*) can be modelled as [Det:(Det|DP) $\gg$ Head:N] ([4]: 354-99, 431-2, 549, 573). Attributive **pre-head AdjPs** and simple **pre-head ING/EN Participials** are ranked similarly as [Mod:(AdjP|V) $\gg$ Head:N] to account for polarity reversals (e.g. *"[trivial problem]$^{(+)}$"*), conflicts (e.g. *"[nasty smile]$^{(-)}$"*), and seemingly contradictory compositions with (?:-) premodifiers (e.g. *"[perfected torture]$^{(-)}$"*). However, mixed sentiment is possible in this construction (e.g. *"[savvy liar]$^{(M)}$"*) ([4]: 444). We rank attributive **pre-head Adverbs** as [Mod:Adv $\gg$ Head:(Adj|Adv)] (e.g. *"[decreasingly happy]$^{(-)}$"*, *"[never graceful(ly)]$^{(-)}$"*) although they too can lead to unresolvable mixed sentiment (e.g. *"[impressively bad(ly)]$^{(M)}$"*) (*idem.* 548, 572-3, 582-5). The pre-head **Negator (Neg)** *"not"*, which is stronger than its head in NPs (e.g. *"[not a scar]$^{(+)}$"*), AdjPs, AdvPs, and PPs, is ranked as [Mod:Neg $\gg$ Head:(N|Adj|Adv|P)] (cf. [7]). In contrast, **pre-head Nouns and Nominals** in NPs are secondary ([Head:N $\gg$ Mod:(N|Nom)]) as seen in polarity conflicts (e.g. *"[family benefit fraud]$^{(-)}$"*, *"[abuse helpline]$^{(+)}$"*) and [¬] head nouns (e.g. *"[risk minimisation]$^{(+)}$"*) (*idem.* 444, 448-9). The genitive subject determiner with the clitic *'s* appears similarly weaker than its head noun or nominal ([Head:(N|Nom) $\gg$ Subj-Det:NP$_{\text{gen}}$]) (e.g. *"[the war's end]$^{(+)}$"*), although polarity conflicts can lead to exceptions: compare *"[the offender's apology]$^{(+)}$"* with *"[the rapist's smile]$^{(-)}$"* (*idem.* 467-83).

Post-head dependents' weights are more variable. In NPs, **post-head AdjPs** generally dominate (e.g. *"[my best friend angry at me]$^{(-)}$"*) as [Comp:AdjP $\gg$ Head:N] (*idem.* 445). **Post-head Participials** dominate their head nouns as [Comp:VP $\gg$ Head:N] (e.g. *"[ugly kids smiling]$^{(+)}$"*, *"[the cysts removed]$^{(+)}$"*) (*idem.* 446), but **post-head VPs** are dominated by their head prepositions ([Head:P $\gg$ Comp:VP]) (e.g. *"[against helping her]$^{(-)}$"*) ([4]: 641). **Post-head PPs** are likewise dominated by their noun, adjective, or adverb heads. The rankings [Head:(N|Adj|Adv) $\gg$ Comp:PP] are thus proposed (e.g. *"[different(ly) from those losers]$^{(+)}$"*, *"[unhappy with success]$^{(-)}$"*, *"[the end of the war]$^{(+)}$"*) ([4]: 446, 543-6). However, exceptions may surface in these constructions, especially in NPs: compare *"[two morons amongst my friends]$^{(-)}$"* with *"[cute kittens near a vicious python]$^{(-)}$"*. Moreover, mixed sentiment may surface (e.g. *"[angry protesters against the war]$^{(M)}$"*). Lastly, we rank **post-head NPs** in PPs as [Head:P $\gg$ Comp:NP] (e.g. *"[against racism]$^{(+)}$"*, *"[with pleasure]$^{(+)}$"*) (*idem.* 635).

In clausal analysis, we treat as the clausal head the predicator (P) which is made of one verb group and compulsory (C)omplements and optional (A)djuncts. The predicator is generally stronger than its complements. We propose that internal complements (Direct Object (O^D), Indirect Object (O^I), Subject Predicative Complement (PC^S), Object Predicative Complement (PC^O), and Oblique (C)omplement) be combined with the predicator before combining the resultant predicate with the predicator's external complements ([4]: 215-8; 236-57). In **Monotransitive Predicates (P-$\mathbf{O^D}$)**, the ranking [Head:P $\gg$ Comp:O^D] models propagation (e.g. *"[failed it]$^{(-)}$"*), polarity conflicts (e.g. *"[spoiled the party]$^{(-)}$"*), and [¬] predicators (e.g. *"[prevent the war]$^{(+)}$"*) (*idem.* 244-8). **Ditransitive Predicates (P-$\mathbf{O^I}$-$\mathbf{O^D}$), (P-$\mathbf{O^D}$-C)** behave in a similar way. Since the monotransitive *"[sent junk]$^{(-)}$"*, pure ditransitive *"[sent me junk]$^{(-)}$"*, and oblique ditransitive *"[sent junk to me]$^{(-)}$"* all share a [-] P-O^D core, we resolve it first before adding an O^I or C to model propagation (e.g. *"[baked a yummy cake for me]$^{(+)}$"*), and polarity conflicts (e.g. *"[brought my friend sad news]$^{(-)}$"*) (*idem.* 244-8). Through the ranking [Head:P $\gg$ Comp:PC^S], typically (N) copular verbs in **Complex Intransitive Predicates (P-$\mathbf{PC^S}$)** can be explained (e.g. *"[seems nice]$^{(+)}$"*) (*idem.* 251-72). **Complex Transitive Predicates (P-$\mathbf{O^D}$-$\mathbf{PC^O}$)** resemble P-PC^S predicates in that the additional direct object does not generally affect the P-PC^S core (e.g. *"[consider (the winner/it/the poison) ideal]$^{(+)}$"*). Hence the ranking [Head:P-PC^O $\gg$ Comp:O^D] (*ibidem*). **(S)ubjects** are ranked as [Head:P $\gg$ Comp:S] (e.g. *"[love can hurt]$^{(-)}$"*, *"[the misery ended]$^{(+)}$"*) (*idem.* 235-43). Note that [¬] NP complements constitute an exception calling for reverse rankings - consider *"[nobody*

PHRASES					
Pre-head			Post-head		
$(\underline{\text{Det}}\text{:(Det}\vert\text{DP)}\vert\text{Subj-Det:NP}_{\text{gen}}{}^{[\neg]}\vert\underline{\text{Mod}}\text{:(Neg}\vert\text{AdjP}\vert\text{V))}$	$\gg$	Head:N	Head:(N$\vert$Nom)	$\ll$	$\underline{\text{Comp}}\text{:(AdjP}\vert\text{VP)}$
$(\underline{\text{Det}}\text{:(Det}\vert\text{DP)}\vert\underline{\text{Mod}}\text{:(Neg}\vert\text{PP}\vert\text{AdvP))}$	$\gg$	Head:Adj	$\underline{\text{Head}}\text{:Adj}$	$\gg$	Comp:PP
$(\underline{\text{Det}}\text{:(Det}\vert\text{DP)}\vert\underline{\text{Mod}}\text{:(Neg}\vert\text{Adv))}$	$\gg$	Head:Adv	$\underline{\text{Head}}\text{:Adv}$	$\gg$	Comp:PP
$\underline{\text{Mod}}\text{:(Neg}\vert\text{AdvP}\vert\text{NP)}$	$\gg$	Head:P	$\underline{\text{Head}}\text{:P}$	$\gg$	Comp:(NP$\vert$VP)
$\text{(Subj-Det:NP}_{\text{gen}}\vert\text{Mod:(N}\vert\text{Nom))}$	$\ll$	$\underline{\text{Head}}\text{:N}$	$\underline{\text{Head}}\text{:N}$	$\gg$	Comp:(NP$\vert$PP)
CLAUSES					
$(\underline{\text{Comp}}\text{:(PC}^{\text{S}}\vert\text{S}^{[\neg]}\vert\text{O}^{\text{D}\,[\neg]}\vert\text{O}^{\text{I}\,[\neg]})\vert\underline{\text{A}}\text{:(AdvP}\vert\text{AdjP}\vert\text{PP)}\vert\underline{\text{Mod}}\text{:Neg)}$	$\gg$	Head:P	$\underline{\text{Head}}\text{:P}$	$\gg$	$\text{Comp:(S}\vert\text{O}^{\text{D}})$
Comp:O^{D}	$\ll$	$\underline{\text{Head}}\text{:P-PC}^{\text{O}}$	$\underline{\text{Head}}\text{:P-O}^{\text{D}}$	$\gg$	$\text{Comp:(O}^{\text{I}}\vert\text{O}^{\text{C}})$

Table 1: Sample Construction Rankings

died]$^{(+)}$", "*[killed nobody]*$^{(+)}$", for example. Hence the rankings [$\underline{\text{Comp}}$:(O$^{\text{D}\,[\neg]}$|S$^{[\neg]}$) ≫ Head:P] for these special cases. Adjuncts are generally stronger than predicators and predicates. The ranking [$\underline{\text{Comp}}$:AdvP ≫ Head:P] for **AdvP Adjuncts**, for example, supports propagation (e.g. "*[he moved it gently]*$^{(+)}$"), and polarity conflicts (e.g. "*[greeted him insincerely]*$^{(-)}$") (*idem.* 224-5, 575, 669, 779-84).

These and other sample rankings are summarised in Table 1.

4 Implementation

The proposed model was implemented as a lexical parsing post-process interpreting the output of a dependency parser[1]. We employ a sentiment lexicon containing manually-compiled atomic core carriers[2] expanded semi-automatically using WordNet 2.1, all tagged with the compositional tags. A morphological unknown carrier guessing module and a missing dependency link repair module are included. Adhering to the proposed compositional processes and constituent rankings at each stage of the analysis, token dependency links and morphosyntactic token tags (e.g. word class, syntactic role, (pre-/post-)head status) are first used to construct individual syntactic phrases (NPs, VPs, AdjPs, AdvPs) and to calculate their internal polarities (**phrasal sentiment**) through stepwise chunking rules which find the rightmost subconstituent in a given phrase and expand it leftwards until a phrasal boundary is hit (see Ex. 2-3). To calculate **clausal** and **sentential sentiment**, the obtained phrasal constituents are then combined (see Ex. 4).

5 Experiments

To estimate the usefulness of a compositional treatment and its impact on standard NLP pipelines, we employ short headlines for sentential compositionality and NPs for phrasal compositionality. Since our implementation is fully lexical, its recall is conditioned by the coverage of the lexicon used. To estimate the (future) impact of larger lexica covering the entire WordNet, the default lexicon (at the time of writing) (DEFAULT_LEX) was expanded with sample carriers from the test data found in WordNet 2.1 (WN_ADD_LEX). Polarity agreement between the gold standards and our output was measured using (*i*) all polarities (*All pol*), and (*ii*) non-neutral polarities only (*Non-ntr pol*). To assess the role of sentiment intensity, results using (*i*) cases of `Any Strength` and (*ii*) those marked as `Strong` in the gold standards are given. The agreement results are shown in Table 2.

Experiment 1: Headlines. The sentences generated by our system were compared against 1000 news headlines in the SemEval-2007 Task #14 data set annotated for polarity (six annotators, *r* .78) ([8]). The SemEval scores [-100, 100] were collapsed into (-100 ≤ (-) < 0; 0 = (N); 0 < (+) ≤ 100) in the `Any Strength` condition, and into (-100 ≤ (-) ≤ -66; 0 = (N); 66 ≤ (+) ≤ 100) for 208 `Strong` cases. The WN_ADD_LEX lexicon contained 97 added carriers.

Experiment 2: NPs. The NPs generated by our system were compared (lax overlap) against 1541 explicit NPs in the customer review data set of 2108 product feature mentions from five home electronics products annotated for polarity (two annotators, *r* unknown) ([3]). The gold standard scores [-3, 3] were converted into (-3 ≤ (-) < 0; 0 < (+) ≤ 3) in the `Any Strength` condition, and into (-3 = (-); 3 = (+)) for 366 `Strong` cases. The WN_ADD_LEX lexicon contained 95 added carriers.

Results and Error Analysis

The *All pol* figures are considerably lower than the corresponding *Non-ntr pol* ones due to the incomplete coverage of the lexica used: a (N) input into the model leads unavoidably to a (N) output and thus to an error. Since mining and tagging new carriers is a task beyond the realms of the model, we focus here on the performance in the *Non-ntr pol* conditions. Mirroring human judgements of high-intensity cases, the implementation performed noticeably better with strong cases. More interesting is the small margin between the two lexica which offers further evidence *pro* compositionality. The errors from the [WN_ADD_LEX *Non-ntr pol* `Any Strength`] condition are analysed in Table 3.

Because the model operates in the middle of the processing pipeline, the errors are classified as *pre-compositional* (i.e. erroneous input) or *post-compositional* (i.e. factors beyond the model). The performance of the model is promising as most errors (ca. 2/3) occurred earlier in the pipeline. Since full compositionality can only be achieved with a clean grammatical analysis, a heavy burden is placed on the TAGGER and PARSER which together caused ca. 28% of the errors. Hence erroneous propagation and partial compositionality due to incorrect POS tags and null dependencies, respectively. Since polarity distinctions between individual word SENSEs (e.g. "*[rip*

[1] Connexor Machinese Syntax 3.8 (`www.connexor.com`)
[2] Kindly provided by Corpora Software (`www.corporasoftware.com`)

	DEFAULT_LEX		WN_ADD_LEX	
Cases	*All pol*	*Non-ntr pol*	*All pol*	*Non-ntr pol*
Headlines				
1000	`Any Strength`			
A	63.0	76.27	65.6	77.36
208	`Strong`			
A	81.73	89.95	86.06	91.33
NPs				
1541	`Any Strength`			
P	72.46	85.45	73.39	85.87
R	97.79	82.93	97.79	83.58
F	83.24	84.17	83.85	84.71
366	`Strong`			
P	79.22	89.10	80.33	89.51
R	98.63	87.70	98.63	88.52
F	87.87	88.40	88.55	89.01

Table 2: Agreement: (A)ccuracy, (P)recision, (R)ecall, and (F)-scores

(a CD)]$^{(N)}$*"* vs. *"[rip into]*$^{(-)}$*"*) can have far-reaching compositional consequences, a sentiment WSD module could reduce ca. 25% of the errors. Resolving neutral ANAPHORic and CO-REFerential expressions could increase recall levels further. The errors also include supraclausal cases NOT yet IMPLemented. However, even in an ideal situation with a clean input, the model would fail to solve many cases (ca. 19%) in which further WORLD knowledge is required. There are cases in which the literal/logical compositional polarity is modulated by phenomena closer to PRAGMatics than lexical semantics such as indirect speech acts (cf. logical positivity vs. implied negativity in *"[X could be better]*$^{(-)}$*"*). Lastly, AMBIGuous cases affording multiple polarity readings are always likely to be present.

	Headlines	**NPs**	**All**	
Pre-compositional errors				
ANAPHOR		13 (6.05)	13	3.23
CO-REF		4 (1.86)	4	0.99
NOT IMPL	8 (4.26)	22 (10.23)	30	7.44
PARSER	13 (6.92)	44 (20.47)	57	14.14
SENSE	58 (30.85)	46 (21.4)	104	25.81
SPELLING		2 (0.93)	2	0.5
TAGGER	24 (12.77)	32 (14.88)	56	13.9
			266	66%
Post-compositional errors				
AMBIG	35 (18.62)	4 (1.86)	39	9.68
PRAGM		21 (9.77)	21	5.21
WORLD	50 (26.6)	27 (12.56)	77	19.11
			137	34%
Total	188	215	403	100%

Table 3: Error distribution

6 Related Work

The proposed model develops further the lexical devices described in the survey of lexical and discoursal contextual valence shifters in ([7]). In ([6]), negation and change phrases were used in a supervised learning algorithm analysing sentential polarities of clinical outcomes. A number of polarity shifters and syntactic dependencies were included as machine learning features in the phrase-level sentiment analyser reported in ([10]). Adjectival appraisal groups comprising a head adjective with optional appraisal premodifiers were used in the sentiment classifier described in ([9]). ([1]) extracted and tagged words with reversal potential expressing a conceptual in-/decrease in magnitude, intensity or quality.

7 Conclusion

We have shown that sentiment exhibits quasi-compositionality in noticeably many areas, and that it is possible to approach sentiment propagation, polarity reversal, and polarity conflict resolution within different linguistic constituent types at different grammatical levels in an analytically and computationally uniform manner by relying on traditional compositional semantics and deep parsing. The results obtained, which are encouraging for a lexical system, point towards a crucial dependency on a wide-coverage lexicon, accurate parsing, and sentiment sense disambiguation in a compositional approach to sentiment analysis.

References

[1] A. Andreevskaia and S. Bergler. Semantic tag extraction using wordnet glosses. In *Proceedings of LREC 2006*, Genoa, 2006.

[2] D. Dowty, R. Wolf, and S. Peters. *Introduction to Montague Semantics.* D. Reidel, Dordrecht, 1981.

[3] M. Hu and B. Liu. Mining and summarizing customer reviews. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD-2004)*, Seattle, 2004.

[4] R. Huddleston and G. K. Pullum. *The Cambridge Grammar of the English Language.* Cambridge University Press, Cambridge, 2002.

[5] S.-M. Kim and E. Hovy. Determining the sentiment of opinions. In *Proceedings of COLING 2004*, Geneva, 2004.

[6] Y. Niu, X. Zhu, J. Li, and G. Hirst. Analysis of polarity information in medical text. In *Proceedings of the American Medical Informatics Association 2005 Annual Symposium (AMIA 2005)*, Washington D.C., 2005.

[7] L. Polanyi and A. Zaenen. Contextual lexical valence shifters. In Y. Qu, J. Shanahan, and J. Wiebe, editors, *Exploring Attitude and Affect in Text: Theories and Applications: Papers from the 2004 Spring Symposium, Technical Report SS-04-07.* AAAI, 2004.

[8] C. Strapparava and R. Mihalcea. Semeval-2007 task 14: Affective text. In *Proceedings of SemEval 2007*, Prague, 2007.

[9] C. Whitelaw, N. Garg, and S. Argamon. Using appraisal taxonomies for sentiment analysis. In *Proceedings of the 2005 ACM CIKM International Conference on Information and Knowledge Management*, Bremen, 2005.

[10] T. Wilson, J. Wiebe, and P. Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of HLT/EMNLP 2005*, Vancouver, 2005.

[11] H. Yu and V. Hatzivassiloglou. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of EMNLP 2003*, Sapporo, 2003.