Generating more-positive and more-negative text

Diana Zaiu Inkpen
School of Information Technology and Engineering
University of Ottawa
Ottawa, ON, Canada, K1N 6N5
diana@site.uottawa.ca

Ol'ga Feiguina and **Graeme Hirst**
Department of Computer Science
University of Toronto
Toronto, ON, Canada, M5S 3G4
{olga,gh}@cs.toronto.edu

Abstract

We present experiments on modifying the semantic orientation of the near-synonyms in a text. We analyze a text into an interlingual representation and a set of attitudinal nuances, with particular focus on its near-synonyms. Then we use our text generator to produce a text with the same meaning but changed semantic orientation (more positive or more negative) by replacing, wherever possible, words with near-synonyms that differ in their expressed attitude.

Near-synonyms and attitudinal nuances

The choice of a word from among a set of near-synonyms that share the same core meaning but vary in their connotations is one of the ways in which a writer controls the nuances of a text. In many cases, the nuances that differentiate near-synonyms relate to expressed attitude and affect. For example, if a writer wants to express a more-favorable view of the appearance of a relatively narrow person, he or she can use the words *slim* or *slender*; if the writer wants to express a less-favorable view, the word *skinny* is available.

This level of attitude expression is distinct from that of the opinions expressed in the text as a whole, and may in fact contradict it. In particular, *euphemism* is the expression of a critical or unpleasant message in relatively positive or favorable terms; *dysphemism* is the converse (Allan & Burridge 1991). Nonetheless, the term *semantic orientation* has been used to describe attitudes at both levels.

Any natural language understanding or generation system must be sensitive to this kind of nuance in text if it is to do its work well. A machine translation system, especially, must recognize such nuances in the source text and *preserve* them in the target text. If the source is, say, polite, angry, or obsequious, then the translation must be too.

Nonetheless, in this paper we look at *changing* the nuances of a text rather than preserving them. We see this primarily as an exercise in the control of nuances in text, and hence a test of a natural language generation system, rather than as a useful application that is an end in itself. That is, any system that purports to accurately preserve nuances should be equally able to change nuances as desired, and render its input in a variety of ways. However, it is possible that a system that can change the nuances of a text could sometimes be helpful — for example, in the customization of texts for users. When generating text that expresses a strong opinion, a negative or positive tone may reflect the speaker's point of view. In this paper, we propose to automatically transform the low-level semantic orientation of a text by choosing near-synonyms accordingly.

In our previous work (Inkpen 2003; Inkpen & Hirst 2001) we automatically acquired a lexical knowledge-base of near-synonym differences (LKB of NS) from the explanatory text of a special dictionary of synonym discrimination, *Choose the Right Word* (hereafter CTRW) (Hayakawa 1994). The main types of distinctions (nuances) that we extracted were: stylistic (for example, *inebriated* is more formal than *drunk*), attitudinal (for example, *skinny* is more pejorative than *slim*), and denotational (for example, *blunder* implies *accident* and *ignorance*, while *error* does not). The computational model we use for representing the meaning of near-synonyms was initially proposed by Edmonds and Hirst (2002).

We enriched the initial LKB of NS with additional information extracted from other sources. Knowledge about the collocational behavior of the near-synonyms was acquired from free text (Inkpen & Hirst 2002). More knowledge about distinctions between near-synonyms was acquired from machine-readable dictionaries: attitudinal distinctions from the *General Inquirer*, and denotational distinctions from word definitions in the *Macquarie Dictionary*. These distinctions were merged with the initial LKB of NS, and inconsistencies were resolved. Our final LKB of NS has 904 clusters containing a total of 5,425 near-synonyms.

The *General Inquirer* (Stone *et al.* 1966) is particularly important in this facet of our work. It is a computational lexicon compiled from several sources, including the Harvard IV-4 dictionary and the Lasswell value dictionary. It contains 11,896 word senses, each tagged with markers that classify the word according to an extensible number of categories. There are markers for words of pleasure, pain, virtue, and vice; markers for words indicating overstatement and understatement; markers for places and locations; etc. The definitions of each word are very brief. Some example entries in GI are presented in Table 1.

The *General Inquirer* category of interest to our work is

CORRECT#1	H4Lvd Positiv Pstv Virtue Ovrst POSAFF Modif 21% adj: Accurate, proper
CORRECT#2	H4Lvd Positiv Pstv Strng Work IAV TRNGAIN SUPV 54% verb: To make right, improve; to point out error (0)
CORRECT#3	H4Lvd Positiv Pstv Virtue Ovrst POSAFF Modif 25% adv: "Correctly" – properly, accurately
CORRECT#4	H4Lvd Virtue TRNGAIN Modif 0% adj: "Corrected" – made right

Table 1: *General Inquirer* entries for the word *correct*.

Positiv/Negativ. (The abbreviations Pstv/Ngtv in Table 1 are earlier versions of Positiv/Negativ.) A positive word corresponds to a favorable attitude; a negative one corresponds to a pejorative attitude. There are 1,915 words marked as Positiv (not including words for *yes*, which is a separate category of 20 entries), and 2,291 words marked as Negativ (not including the separate category *no* in the sense of refusal). An attitudinal distinction was asserted in our LKB of NS for each near-synonym in CTRW that was marked Positiv or Negativ in GI.

In this paper, we focus on the attitudinal distinctions stored into our LKB of NS, acquired from CTRW and GI. For our near-synonyms, we extracted 1,519 attitudinal distinctions from GI, and 384 from CTRW. The information acquired from the two sources was merged and conflicts were resolved through a voting scheme. After merging, we were left with 1,709 attitudinal distinctions in our LKB of NS. The rest of the near-synonyms are considered neutral by default.

Related work

There is much recent work on the classification of text (at the document level or at the sentence level) as objective or subjective (Riloff & Wiebe 2003), and the classification of subjective text as positive or negative (Turney 2002; Pang, Lee, & Vaithyanathan 2002; Yu & Hatzivassiloglou 2003). Work on generation using pragmatic nuances, including the attitude of the speaker and of the hearer was presented by Hovy (1990). Elhadad (1997) presented work on unification-based constraints for lexical choice in generation. Similarly, our generator uses collocations to constrain the lexical choice, but it also includes the possibility of expressing lexical nuances.

Our work in this paper has a different focus, on the analysis of subjective text, extracting its lexical nuances (including attitude), and generating a text with the same meaning but a new semantic orientation. This is, in effect, translating from English to English via an interlingual representation, changing the semantic orientation before the generation phase.

Estimating the relative semantic orientation of text

We extracted paragraphs from the British National Corpus (BNC) that contain at least three of our set of near-synonyms. We chose to use paragraphs because we believe that the change in orientation will be more noticeable than at the sentence level, and more localized than at the document level (because we cannot be sure that the semantic orientation does not change from paragraph to paragraph in the BNC).

We did not classify the complete texts according to their semantic orientation. We only estimated, semi-automatically, the orientation of each selected paragraph from the semantic orientation of its words. We labeled as many words (except stopwords) as we could as positive, negative, or neutral as follows. First, we checked whether the word is a near-synonym in our LKB of NS. If so, we consulted the LKB regarding the attitude of the near-synonym. We did sense disambiguation as described in the next section. We consulted the GI for the attitude of all the other words. The sense disambiguation mechanism for this part is also described in the following section. A majority vote gave us an estimate of the attitude of the paragraph: Favorable, Pejorative, or Neutral. We declared a paragraph to be Neutral (not subjective) if fewer than three pejorative or favorable words were discovered.

There are several problems with this approach, related to the fact that we look at individual words and ignore longer expressions. First, neighboring words can change the attitude of a word (e.g. *not good* is negative while *good* is positive). Second, words may have different attitudes when they are used as part of an expression or collocation (e.g., *out to lunch* is negative while the individual words are neutral or positive) (Baron & Hirst 2004). Lastly, the author may be employing irony or sarcasm, which is not detected by our method. Another limitation is that if the information in the LKB of NS for a word was acquired from CTRW the near-synonyms are classified as favorable, pejorative, or neutral only in comparison to other near-synonyms in their cluster; that is, the classification is relative. For example, *mistake* is Favorable in the LKB of NS because it's better than *blunder*, but the word *mistake* itself is not very positive. Despite these problems, because we look at the words in a paragraph and take a majority vote, we can determine the probable correct semantic orientation of the paragraph.

We also experimented with paragraphs from Epinions (*www.epinions.com*), a Web site where users review and rate books, movies, music, and various products and services. The reviews are typically several paragraphs long, and are accompanied by a rating on a scale of one to five stars. If a user rates an item with four or five stars, we can assume that the text of the associated review is positive. If the rating is one or two stars, we can assume that the text is negative.

Word sense disambiguation

When looking up the attitude of a word in our LKB or in GI, we needed to first disambiguate it, because the nuances of a word may depend on the sense in which it is being used.

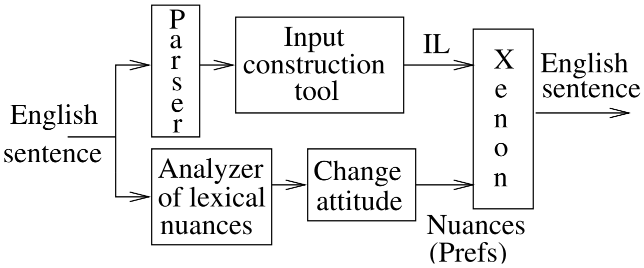

Figure 1: The architecture of the system.

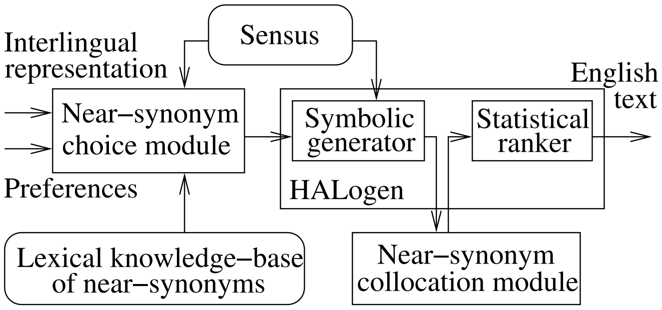

Figure 2: The architecture of Xenon.

Since the BNC text is PoS-tagged, we could rule out senses with a different part of speech. After that, when looking up words in the GI, we just took the most frequent sense. In our LKB of NS, different senses of a near-synonym can belong to different clusters of near-synonyms. We also had situations when a word in the paragraph, considered a potential near-synonym, was used in a sense that was not in the LKB of NS. For example, the word *blue* is in the LKB in the sense of *sad*, but not in the sense of a color. So, we had to consider every cluster and decide whether it is the right sense. We attempted to do this by checking whether the intersection of the paragraph and the text of the CTRW entry for this cluster (both considered as bags of words, with stopwords removed) was empty or not. But this did not work well. So we completed the near-synonym sense disambiguation in a semi-automatic manner, by hand-correcting the wrong decisions. In later work, we hope to improve the sense disambiguation module by using semantic relatedness instead of a simple intersection. Disambiguation of near-synonym senses is also used in the analysis module that will be presented in the next section.

Analysis

Figure 1 presents the global architecture of our system. Each sentence of the paragraph was parsed with Charniak's parser (Charniak 2000), and we applied an input construction tool,[1] which produces a shallow interlingual representation (IL) from each parse tree. This will be described in the next section. We then substituted a meta-concept — a disjunction of the near-synonyms of the initial near-synonym — for each near-synonym in the interlingual representation.

Generation

After analysis of the input, the resulting interlingual representation and the set of lexical nuances are input to the generator module, which is named Xenon (see figure 1). The set of lexical nuances become preferences to be satisfied by Xenon. But before the lexical nuances are passed to Xenon, those relating to attitude may be modified as desired by the user.

[1]Thanks to Irene Langkilde-Geary for making the input construction tool available. She built this tool to produce inputs for the evaluation of the HALogen generator.

Xenon (Inkpen & Hirst 2003) is our natural language generation system, capable of distinguishing between near-synonyms in generation. Xenon integrates a new near-synonym choice module and a near-synonym collocation module with the HALogen sentence realization system (Langkilde & Knight 1998; Langkilde-Geary 2002). HALogen is a broad-coverage general-purpose natural language sentence generation system that combines symbolic rules with a language model derived from large text corpora. For a given input, it generates all the possible English sentences into a compact forest representation and then ranks the sentences according to its language model, in order to choose the most likely sentence as output. Xenon extends this, using the LKB of NS and a set of desired nuances to possibly override the choice that HALogen would otherwise make.

The IL input to Xenon, like the input to HALogen, is expressed in an interlingua developed at the Information Sciences Institute, University of Southern California (ISI). This language contains a specified set of 40 roles, and the fillers of the roles can be words, concepts from Sensus (Knight & Luk 1994), or complex representations (Langkilde-Geary 2002). Xenon extends this representation language by adding meta-concepts that correspond to the core denotation of the clusters of near-synonyms (disjunctions of all the near-synonyms in a cluster).

Figure 2 presents the architecture of Xenon. The input is a semantic representation and a set of preferences to be satisfied. The final output is a set of sentences and their scores. An example of input and output is shown in Figure 3. The first sentence (the highest-ranked) is considered to be the solution. In this example, *fib* was chosen from the cluster *lie, falsehood, fib, prevarication, rationalization, untruth* to represent the meta-concept `generic_lie_n`.

The near-synonym choice module chooses the near-synonym from each cluster that best matches the input preferences. The preferences, as well as the distinctions between near-synonyms stored in the LKB of NS, are of three types. Stylistic preferences express a certain formality, force, or concreteness level and have the form: (*strength stylistic-feature*), for example `(low formality)`. Denotational preferences connote a particular concept or configuration of concepts and have the form: (*indirectness peripheral-concept*), where *indirectness* takes one of the values `suggest, imply, denote`.

```
Input:
(A9 / tell
      :agent (V9 / boy)
      :object (O9 / generic_lie_n))
Input preferences:
      ((DISFAVOR :AGENT)
        (LOW FORMALITY)
        (DENOTE (C1 / TRIVIAL)))
Output:
The boy told fibs.  −40.8177
Boy told fibs.  −42.3818
Boys told fibs.  −42.7857
```

Figure 3: Example of input and output of Xenon.

An example is: `(imply (C / assessment :MOD (OR ignorant uninformed)))`. The peripheral concepts are expressed in the ISI interlingua. And attitudinal preferences, which are the ones that are of special interest here, express a favorable, neutral, or pejorative attitude and have the form: (*stance entity*), where *stance* takes one of the values `favor, neutral, disfavor`. An example is: `(disfavor :agent)`.

The near-synonym collocations module ensures that the text generated does not contain unacceptable collocations. Near-synonyms that would violate collocational constraints are assigned lower weights, so that they will not be chosen by later processes. Possible collocations are detected in the forest representation that is output by HALogen's symbolic generator, the weights are decreased as needed, and the modified forest representation is input to HALogen's statistical ranker to finish the generation. The near-synonym collocations module is important in generating text with different semantic orientations, because by simply replacing a negative near-synonym with a positive one we might violate collocational constraints.

Experiments

We ran Xenon on the IL representations that resulted from the analysis, as described above, of each of the paragraphs that we selected from the BNC and from Epinions. Figure 4 shows an example of paragraph from Epinions, part of a negative review with an accompanying rating of two stars.

In the first two experiments, the set of preferences contained only one element, an attitudinal preference. We generated paragraphs with positive orientation, using the preference `(favor :agent)`. The paragraph generated for our example paragraph is presented in Figure 5.[2] We did not consider the semantic orientation of the initial paragraph; we simply generated positive or negative text. If the original paragraph was negative, the generated text is expected to be more positive; if the original paragraph was already positive, the generated text should be the same or slightly more positive. Similarly, we generated negative paragraphs with the input preference `(disfavor :agent)`. The paragraph generated for our example is presented in Figure 6. In our example, the initial paragraph was relatively negative (two stars), but we expect it to become even more negative (corresponding to a one-star rating).

During my trip to Aruba a few years back, my boyfriend and I rented a car and drove around the island. While looking for the natural bridge we found a tourist spot called the "Tunnel of Love". Essentially, it's a big cave that you pay to walk through. We thought it would be nice to experience some of the natural beauty Aruba had to **offer**. I had just hurt my ankle in a jetskiing incident and asked if I would be **able** to walk through with a gimpy leg. The women at the entrance told me I should have no **trouble**. The cave starts out as a pretty large enclosed space, with some external light and quite **easy** to maneuver. As you progress inward and downward, the space gets narrower, darker and more difficult to walk through. At approximately halfway through I literally had to hunch over to pass through. That's when the funny **smell**, strange noises and incredible heat kicked in and my light switched off! So here we are in the pitch blackness of a hot and **humid cave**. When our light flickered on for a few moments, the rays of illumination happened to pass over our fellow cave dwellers — a colony of bats. I later learned that the funny **smell** is bat waste! This is where I almost had a coronary and picked up the pace forward. When we finally reached the end I found out that instead of walking out, you climb out! With only one good leg and the other to use only as a support, I had a lot of **trouble** getting out. Luckily the **smell** of bat urine, got me moving. As a reasonable healthy bodied person I was slightly inconvenienced but **elderly** and sickly people who visit the Tunnel of Love may have some **serious** issues with this tourist spot.

Figure 4: An example of original paragraph.

We also experimented with a set of preferences that preserves the original nuances of near-synonyms in the text, and adds `(favor :agent)` or `(disfavor :agent)`. The attitudinal preference is given a higher importance than the rest of the preferences in order to increase the change in semantic orientation as much as possible. In these experiments, we expected a near-synonym in the paragraph to change only if there was another near-synonym in the same cluster with the desired orientation and with lexical nuances that are not incompatible with the initial nuances. The resulting positive paragraph is very similar to the one presented in Figure 5, with small differences; but the word *aroma* was replaced by the word *smell*, which was also used in the original paragraph. This is what we expected to obtain by preserving lexical nuances; the word *aroma* is more positive, but it introduces the nuance of a very pleasant smell, which is not the case in this text. The negative paragraph is very similar to that presented in Figure 6, with the difference that more words were chosen as in the original paragraph: *stink* was

[2] In order to focus on the lexical-choice issues, rather than choice of syntactic structures and the limitations of HALogen and Xenon, figures5 and 6 do not show the actual output, but rather the crucial lexical choices substituted back into the original paragraph. The actual output is very close to the text we showed, with a few small exceptions.

During my trip to Aruba a few years back, my boyfriend and I rented a car and drove around the island. While looking for the natural bridge we found a tourist spot called the "Tunnel of Love". Essentially, it's a big cave that you pay to walk through. We thought it would be nice to experience some of the natural beauty Aruba had to **offer**. I had just hurt my ankle in a jetskiing incident and asked if I would be **able** to walk through with a gimpy leg. The women at the entrance told me I should have no **exertion**. The cave starts out as a pretty large enclosed space, with some external light and quite **easy** to maneuver. As you progress inward and downward, the space gets narrower, darker and more difficult to walk through. At approximately halfway through I literally had to hunch over to pass through. That's when the funny **aroma**, strange noises and incredible heat kicked in and my light switched off! So here we are in the pitch blackness of a hot and **humid tunnel**. When our light flickered on for a few moments, the rays of illumination happened to pass over our fellow cave dwellers — a colony of bats. I later learned that the funny **aroma** is bat waste! This is where I almost had a coronary and picked up the pace forward. When we finally reached the end I found out that instead of walking out, you climb out! With only one good leg and the other to use only as a support, I had a lot of **exertion** getting out. Luckily the **odor** of bat urine, got me moving. As a reasonable healthy bodied person I was slightly inconvenienced but **elderly** and sickly people who visit the Tunnel of Love may have some **serious** issues with this tourist spot.

Figure 5: Generated positive text.

During my trip to Aruba a few years back, my boyfriend and I rented a car and drove around the island. While looking for the natural bridge we found a tourist spot called the "Tunnel of Love". Essentially, it's a big cave that you pay to walk through. We thought it would be nice to experience some of the natural beauty Aruba had to **propose**. I had just hurt my ankle in a jetskiing incident and asked if I would be **good** to walk through with a gimpy leg. The women at the entrance told me I should have no **trouble**. The cave starts out as a pretty large enclosed space, with some external light and quite **simplistic** to maneuver. As you progress inward and downward, the space gets narrower, darker and more difficult to walk through. At approximately halfway through I literally had to hunch over to pass through. That's when the funny **stink**, strange noises and incredible heat kicked in and my light switched off! So here we are in the pitch blackness of a hot and **oppressive tunnel**. When our light flickered on for a few moments, the rays of illumination happened to pass over our fellow cave dwellers — a colony of bats. I later learned that the funny **stink** is bat waste! This is where I almost had a coronary and picked up the pace forward. When we finally reached the end I found out that instead of walking out, you climb out! With only one good leg and the other to use only as a support, I had a lot of **trouble** getting out. Luckily the **stink** of bat urine, got me moving. As a reasonable healthy bodied person I was slightly inconvenienced but **old** and sickly people who visit the Tunnel of Love may have some **grave** issues with this tourist spot.

Figure 6: Generated negative text.

replaced by *smell*, *propose* by *offer*, and *good* by *able*.

Sometimes, when we expect a specific word to be chosen because of its semantic orientation, another word might be chosen instead by HALogen's statistical ranker, as it tends to favor frequent words. Also, notice that the choice of near-synonyms can be sometimes infelicitous. For example, the choice of *good to* instead of *able to* in Figure 6 makes the sentence sound odd. The word *good* was included in the near-synonym cluster of *able* by the lexicographers who wrote CTRW, but it was intended as a modifier (e.g., *a good teacher*). We would have expected HALogen's trigram language module to prefer *able to*, since it favours good collocations with function words. Xenon's collocations module favours good collocations between near-synonyms and content words, but the coverage of our collocational knowledge-base is limited.

Evaluation

Our evaluation is at a preliminary stage. We conducted an evaluation that involved human judges comparing paragraphs in terms of attitude. We selected three paragraphs from the BNC, and generated the positive and negative versions of each of them. Hence, we had six pairs of paragraphs for which we wanted the judges to decide, for two paragraphs P_1 and P_2 whether:

- P_1 is more positive (less negative) than P_2;
- P_1 is more negative (less positive) than P_2; or
- P_1 and P_2 are equally positive/negative.

We had six judges in total. Each pair of original paragraph and generated paragraph was presented to three judges. No judge saw the original paragraph more than once. The order of pairs and paragraphs in pairs was randomized. The results are presented in Table 2, where $P_i < P_j$ means P_i was judged to be more negative (or less positive) than P_j. In an ideal case, if the generated texts had the intended semantic orientation, the expected judgments would be $<$ for the upper half of the table (the rows numbered 1, 2, and 3) because the original paragraph should be less positive than the positive-generated one, and $>$ for the lower part (the rows numbered 4, 5, and 6), because the original paragraph should be more positive than the negative-generated one. For the positive texts, the judgments were closer to the expectations than for the negative texts, for which there were too many cases when the judgment was opposite to the expected one. The '=' answers are sometimes consistent with our expectations. This depends on how negative or positive the initial text was. For example, if a text is positive to start with, the more-positive version that we generated might be only slightly more positive, in which case both $>$ and = should be expected judgments. We have yet to include this factor in our results.

The task proved to be difficult for the judges, who were native speakers of English but had no knowledge of com-

No.	Paragraph pair	J_1	J_2	J_3
1	P_1 and P_1-positive	=	$<$	$>$
2	P_2 and P_2-positive	$<$	$<$	=
3	P_3 and P_3-positive	=	$>$	=
No.	**Paragraph pair**	J_4	J_5	J_6
4	P_1 and P_1-negative	=	$>$	=
5	P_2 and P_2-negative	$<$	=	$<$
6	P_3 and P_3-negative	$<$	$>$	$<$

Table 2: Evaluation of the generated texts by judges J_1 to J_6. $P_i < P_j$ means P_i was judged to be more negative than P_j, $P_i > P_j$ means that P_i was judged to be more positive than P_j, and $P_i = P_j$ means they were judged to be equally positive or negative.

putational linguistics. The instructions were short, relying on the judges' intuition of what a positive or negative text is. The results of the preliminary evaluation are not very conclusive; the scale of the experiment was too small. We plan to redo it with more paragraphs from Epinions, which have the advantage that we know their initial semantic orientation. We also plan to choose paragraphs with a higher number of near-synonyms that have positive and negative alternatives in their clusters, and to use more paragraphs and more, better-instructed judges in order to conduct a more-conclusive evaluation.

Conclusion

The work reported here is a pilot for research that is presently in progress, and we have yet only a small amount of data and analysis.

In later work, we need to increase the coverage of our LKB of NS, which is large but not large enough. We need to know what the near-synonyms of all the content words in a paragraph are. It would be useful to know the near-synonyms at least of all the words from GI for which the semantic orientation is known. We could acquire these near-synonyms from corpora. Steps in this direction were taken by Glickman & Dagan (2003) and Lin *et al.* (2003). We also need to acquire attitudinal nuances for the newly-added near-synonyms.

Once the coverage of our LKB of NS is increased, we plan to use other texts, such as movie reviews for which the semantic orientation is known, to generate the same texts with changed orientation. We can also use texts whose semantic orientation is determined by one of the existing classifiers mentioned in the related work section.

Acknowledgments

We thank Irene Langkilde-Geary for making HALogen and the input construction tool available. Our work is financially supported by the Natural Sciences and Engineering Research Council of Canada, the University of Toronto, and the University of Ottawa.

References

Allan, K., and Burridge, K., eds. 1991. *Euphemism and Dysphemism*. Oxford University Press.

Baron, F., and Hirst, G. 2004. Collocations as cues to semantic orientation. Unpublished manuscript.

Charniak, E. 2000. A maximum-entropy-inspired parser. In *Proceedings of the 1st Conference of the North American Chapter of the Association for Computational Linguistics and the 6th Conference on Applied Natural Language Processing (NAACL-ANLP 2000)*.

Edmonds, P., and Hirst, G. 2002. Near-synonymy and lexical choice. *Computational Linguistics* 28 (2):105–145.

Elhadad, M.; McKeown, K.; and Robin, J. 1997. Floating constraints in lexical choice. *Computational Linguistics* 22 (4):1–44.

Glickman, O., and Dagan, I. 2003. Identifying lexical paraphrases from a single corpus: A case study for verbs. In *Proceedings of the International Conference RANLP-2003 (Recent Advances in Natural Language Processing)*.

Hayakawa, S. I., ed. 1994. *Choose the Right Word*. Second Edition, revised by Eugene Ehrlich. HarperCollins Publishers.

Hovy, E. 1990. Pragmatics and language generation. *Artificial Intelligence* 43:153–197.

Inkpen, D. Z., and Hirst, G. 2001. Building a lexical knowledge-base of near-synonym differences. In *Proceedings of the Workshop on WordNet and Other Lexical Resources, Second Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL 2001)*, 47–52.

Inkpen, D. Z., and Hirst, G. 2002. Acquiring collocations for lexical choice between near-synonyms. In *Proceedings of the Workshop on Unsupervised Lexical Acquisition, 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002)*, 67–76.

Inkpen, D. Z., and Hirst, G. 2003. Near-synonym choice in natural language generation. In *Proceedings of the International Conference RANLP-2003 (Recent Advances in Natural Language Processing)*.

Inkpen, D. 2003. *Building a Lexical Knowledge-Base of Near-Synonym Differences*. Ph.D. Dissertation, University of Toronto.

Knight, K., and Luk, S. 1994. Building a large knowledge base for machine translation. In *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94)*.

Langkilde, I., and Knight, K. 1998. Generation that exploits corpus-based statistical knowledge. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics joint with 17th International Conference on Computational Linguistics (ACL-COLING'98)*, 704–712.

Langkilde-Geary, I. 2002. *A Foundation for a General-Purpose Natural Language Generation: Sentence Realization Using Probabilistic Models of Language*. Ph.D. Dissertation, University of Southern California.

Lin, D.; Zhao, S.; Qin, L.; and Zhou, M. 2003. Identifying synonyms among distributionally similar words. In *Proceedings of the Eighteenth Joint International Conference on Artificial Intelligence (IJCAI-03)*.

Pang, B.; Lee, L.; and Vaithyanathan, S. 2002. Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP-02)*.

Riloff, E., and Wiebe, J. 2003. Learning extraction patterns for subjective expressions. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP-03)*.

Stone, P. J.; Dunphy, D. C.; Smith, M. S.; Ogilvie, D. M.; and associates. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. The MIT Press.

Turney, P. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002)*.

Yu, H., and Hatzivassiloglou, V. 2003. Towards answering opinion questons: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP-03)*.