

基于現代漢語語法信息詞典的詞語情感評價研究¹

Research on Lexical Emotional Evaluation Based on the Grammatical Knowledge-Base of Contemporary Chinese

王治敏*、朱學鋒*、俞士汶*

Zhimin Wang, Xuefeng Zhu and Shiwen Yu

摘要

本文將情感評價屬性特徵納入現代漢語語法信息詞典的詞語屬性描述體系，基于人民日報基本標注語料庫，探討以定性和定量相結合的方式對漢語詞語的情感標注進行研究。根據真實文本實例的統計、歸納，對詞典中詞語的情感傾向加以描述，然后在詞典中形式化。詞語的情感評價屬性的計算處理對文本過濾、信息抽取、網頁評價等有重要的參考價值。

關鍵字：詞語情感評價、語義韻律、搭配規律

Abstract

This paper introduces the attributes of emotional evaluation in the Grammatical Knowledge-base of Contemporary Chinese. Lexical emotion tagging is studied by means of both qualitative and quantitative approaches. Based on the statistical results from the People's Daily tagging corpus, lexical emotional trends are described and formulated in our Knowledge-base. Lastly, we also discuss potential applications of emotion tagging in related tasks such as text filtering, information retrieval and web page evaluation.

¹ 相關研究得到中國國家 973 項目(2004CB318102)和國家 863 計劃(2001AA114210, 2002AA117010)的支持。

* 北京大學計算語言學研究所，100871 中國
Institute of Computational Linguistics, Peking University, 100871 China
E-mail: {wangzm, yusw}@pku.edu.cn

Keywords: Evaluation of Lexical Emotion, Semantic Prosody, Collocation Regulation

1. 引言

隨著信息處理領域“信息檢索、文本過濾、自動文摘、網頁評價”等技術的不斷發展，研究者開始利用詞語所表現出來的情感屬性來提高實用系統的智能水準，出現了以網頁評價、產品評價為目標的情感分析研究。例如：[T'sou *et al.* 2005] 運用詞語的情感屬性來衡量名人的公眾看法；[Sanjiv *et al.* 2001] 從股票信息板塊獲取股票市場的評價；[魯紅英 2003]關於名人網頁的相關度評價研究等等。因此詞語表現出來的正面、負面的情感評價屬性特徵越來越受到學者們的關注。

經研究，我們發現詞語蘊涵的情感屬性對其句子中共現的詞語有很大的限制，其共現詞語往往也要求具有統一的情感傾向。例如：以“潰逃”為例，“潰逃”是個貶義詞，當它進入句子中與其共現的主語成分大都是含有貶義的壞人。例如：敵軍～、匪軍～、反動派～、土匪～、壞蛋～、罪犯～、走私犯～。與其共現的狀語成分也表示貶義的含義。例如：倉惶～、狼狽～。也有互為共現的詞語表現出不一致的情感傾向。例如：“擺脫”不是貶義詞，但通常與表示消極情感傾向的詞語共現，如：～困難、～困境、～貧困、～不發達狀態、～羞恥和孤獨、～危機、～老套套、～束縛、～危險、～制裁、～困擾。雖然後面所帶的詞語都是表示消極、負面的，但是整個句子卻表現一種積極、肯定的情感傾向。由此“擺脫”也帶上了積極、正面的色彩。

母語是漢語的中國人也許在毫無察覺的情況下下意識地運用詞語的情感色彩，而對外國留學生或計算機則需要學習才會理解。這些規律如果能夠從真實語料庫中提取，然後對這些規律進行定量的分析，形式化到知識庫中，無論對中文信息處理還是對外漢語教學都是很有價值的。

詞語情感評價在受限領域已經有了一些研究和探索。例如：[蘇玉梅 2004]提出了汽車領域網頁褒貶相關度評價模型。該模型對網頁實體進行褒貶態度評價，其中包含了一系列評價要素，如褒貶結構、領域標準等關鍵方法。該模組被應用到天網知名度系統，得到了很好的試驗效果。

對於漢語通用領域的詞語情感色彩評價研究，目前還沒有人利用大規模文本來做這樣的事。北京大學計算語言所長期致力於中文信息基礎資源的研究和開發。其重要的研究成果《現代漢語語法信息詞典》（簡稱《語法信息詞典》）和人民日報標注語料庫為詞語情感評價研究提供了非常好的基礎資源。語法信息詞典共計收詞 73000 多條²。該詞典在中文信息處理的自動分詞、詞語標注、機器翻譯、信息提取、信息檢索、概念詞典建設等方面發揮了較大的作用[俞士汶 2003]。

這兩年北大計算語言所又得到了 863 項目的支援。預期達到的目標是在突破關鍵技術基礎上，研製符合奧運多語言信息服務需求的大規模通用基礎資源，並建設綜合型語

² 這是《語法信息詞典》2003 年的詞條數量，截止到 2004 年 12 月，詞條數量已經增至 8 萬條。

言知識庫。語法信息詞典的擴充作為其中一個子任務，要求在 2004 年年底增至到 8 萬詞條，利用這次詞典擴充的機會，擬將情感評價特徵納入語法信息詞典的屬性描述體系。

2. 情感評價屬性的界定

關於詞語的情感評價，中國語言學界稱之為詞語的感情色彩（詞的褒貶）。相關的詞彙學著作有過論述[符淮青 1985; 劉叔新 1990]，但他們只是簡單地說明詞語感情色彩的定義和枚舉相應的例子，語法信息詞典的情感評價屬性描述和傳統語言學的詞語感情色彩研究有一定區別。為了能夠和語言學的感情色彩相區別，我們在以後的描述中使用“情感評價”這一術語。

北大現代漢語教材[1997]指出：感情色彩指詞義所附帶的表示褒貶態度的色彩。詞的感情色彩同詞的意義關係密切，詞義對客觀事物有肯定評價的，一般有褒揚色彩。如：英雄、勇士、悅耳、富饒…，詞義對客觀事物有否定評價的，一般有貶斥的感情色彩。如：奸賊、賭徒、陰險、平庸…。從上述的定義可以看出詞語的褒貶色彩是一種表示程度很高的情感評價，這部分詞語在實際語言當中比例並不是很多，大多數詞語雖然無法說出它們的褒貶，但在語言環境中可以表現出積極或消極（正面或負面）的情感傾向。這些詞語在句子中表現出來的情感傾向和詞語共現所表現出來的搭配規律很少有人關注，而這部分信息對於文本過濾，信息安全、網頁評價有很重要的應用價值。我們以“陷入”為例，“陷入”在感情色彩方面並無貶義色彩，但是在具體的語言環境中往往要求其共現詞語含負面、消極的因素。由此“陷入”也表現出消極、負面的情感傾向，這點我們在下面的例句中得到了驗證。陷入(～惡性循環狀態、～苦悶邊緣、～困境、～金融危機、～金融麻煩、～被動、～僵局、～困擾、～危險、～窘境、～被壓迫民族的地位、～惡性循環、～一連串的追殺之中、～彷徨、～庸人的狹小圈子裏、～了一個雷區、～衰退、～癱瘓、～混亂、～了地獄、～低潮、～嚴重的分裂、～孤立狀態、～茫然～、一個怪圈，一個誤區、～低谷、～恐慌、～重重包圍、～沉思、～凝重的沉思)。

“陷入”的使用頻率很高，僅僅兩個月的人民日報就出現了 33 次，表現出負面信息的句子有 32 例可以得到確認，只有一例負面評價不好確認，如：～沉思，“沉思”在現漢中是“深思”之義，表示一種中性含義，例句原文如下：“吳書記眉頭漸漸鎖緊，陷入了沉思。”回到原文就可發現“沉思”所在的語境所表現出來也是一種愁苦的樣子，也應該是負面含義。同時還找到了一句“凝重的沉思”，也是表示一種不好的心緒。以此我們可以得出一個結論，“陷入”通常表示消極、負面的評價。與“陷入”同義的還有“跌入、陷沒、陷落”，它們在句子中的共現詞語也同樣表示消極、負面的感情色彩。動詞短語的中心語動詞“陷入”和其后面所共現的名詞的情感標注趨於一致，給定了“陷入”和其后面共現詞語的情感屬性，機器就可以判定關鍵詞所在句子所描述事件的好壞。因此語法信息詞典對詞語的描述並不限于褒貶的評價，同時還要對詞語進行正面、負面等不同程度的評價。

語言是動態的，靜態詞庫裏的詞語在實際的語言應用中會受到其搭配詞語的影響，一個本來沒有多少褒貶意義的詞語在進入句子框架后可能會表現出強烈的情感傾向。這

方面的研究，國外已經開始得到重視。他們稱之為語義韻律。[Partington 1996]將語義韻律定義為“超越單個詞語界限的色彩意義的擴展”(Semantic prosody refers to the spreading of connotational coloring beyond single word boundaries)，語義韻律由音韻學上的“韻律”概念轉化而來，音韻學上的韻律現象是用來概括語音研究中切分成分在語流中具有超切分特徵。音韻學中的同化、異化、連讀都屬於韻律的研究範疇。受這種現象的啓發，[Sinclair 1987, 1991]把韻律加以推廣，認為這種現象同樣存在語言的詞彙層面，同時利用大規模的語料庫資源對詞語的韻律進行了搭配意義方面的研究。

綜合國內外的研究成果，詞語的情感評價應該從兩個層面來考慮。一方面應該是靜態詞彙層面的研究，即詞語在靜態詞庫中所表現出來的褒義、正面、負面、貶義等情感屬性，這些屬性可以直接在語法信息詞典中描述。另一方面詞語的情感信息在進入句子框架下會發現情感偏移現象，即詞語評價屬性的動態句法研究。詞語情感評價的動態研究將會有助於發現詞語之間的動態搭配規律，同時對於描述語言規律，探求語言認知心理有很好的啓示。

3. 語法信息詞典的評價類別

漢語中的詞語具有情感評價的詞語並不限于名詞、形容詞，其他詞類的詞語也有類似的情感評價傾向。語法信息詞典中可以加上情感色彩描述的有 12 類。具體詞類如下：

名詞(n)

- (1) 春光 慈父 牛市 英雄 勇士 勞模
- (2) 癌細胞 愛滋病毒 悲歌 弊端 殘骸 慘禍 熊市 黑窩 痞子

動詞(v)

- (1) 鍛煉 發明 發揚 防止 奉獻 改善 感謝 鼓勵 灌溉 激發 獎勵
- (2) 暴虐 爆發 爆炸 貶低 瀕臨 殘殺 挫傷 顛覆 妒忌 訛詐 妨礙

形容詞(a)

- (1) 美麗 聰穎 恭敬 燦爛 光滑 單純
- (2) 傲慢 暴虐 悲淒 憋悶 沉痛 惆悵

狀態詞(z)

- (1) 碧油油 光燦燦 甜絲絲 水汪汪 美滋滋 熱熱鬧鬧
- (2) 悲慘慘 癡呆呆 病歪歪 惡狠狠 瘋顛顛 邈裏邈邊

區別詞(b)

- (1) 錦繡(～中華～風光～前程) 稀世(～珍寶～珍品～之寶～杰作)
- (2) 填鴨式(～教學)劣質(～食品) 偽劣(～產品) 違禁(～物品)冒牌(～貨～開發商)

副詞(d)

- (1) 乘興(～追擊) 穩步(～發展～推進～提高～反彈～增長～上揚～升值)豁然(～開

朗) 竭誠(～服務) 闊步(～前進) 銳意(～進取) 捨身(～拼搏～救人～爲國)

(2) 大肆(～燒殺～翻供～索要收取錢物～進行分裂祖國的活動～進行盜竊)

公然(～對鄰國巴基斯坦進行威脅～在巴大肆進行國家恐怖主義活動～敲詐開出天價～拒絕～縱容其進行製造“兩個中國”，～造假，偽造歷史文獻)

乘機(～搗亂～鑽洞而入～發洩～甩包袱) 遲遲(～未察覺～不表態)

不巧(～中國選手來日后多人患感冒～東京的櫻花開得早)

不慎(～掉到地上摔成重傷～買了偽劣產品)成心(～把泡菜醃老了)

嘆詞(e)

(1) 哈哈(～，我鯊膽有救了，膽子還可大一點) 好傢伙(～，當時我差點暈過去)

呵呵(～，就來就來)

(2) 嗚呼(～！她不死于盜匪之手而死于親人之口，真是天下第一等大悲了！)

(～，如果人心與人心之間都掛著防賊的鎖，那單位還有生氣嗎？)

成語(i)

(1) 百廢俱興 比翼雙飛 別具匠心 彬彬有禮 博古通今 赤膽忠心 寵辱不驚

(2) 暗箭傷人 黯然失色 稗官野史 笨嘴拙舌 遍體鱗傷 不擇手段 稱王稱霸

慣用語(l)

(1) 飽眼福 爆冷門 賣力氣 鳴不平 鬧新房 上檔次

(2) 不入流 吃獨食 吃啞巴虧 穿小鞋 發善心 鼓倒掌 悶葫蘆 上賊船 背包袱

簡稱(j)

(1) 五講四美 雙擁 兩彈一星

(2) 死緩 老弱病殘 假冒偽劣 危舊房

代詞(r)

(1) 您 足下

(2) 鄙人 吾輩

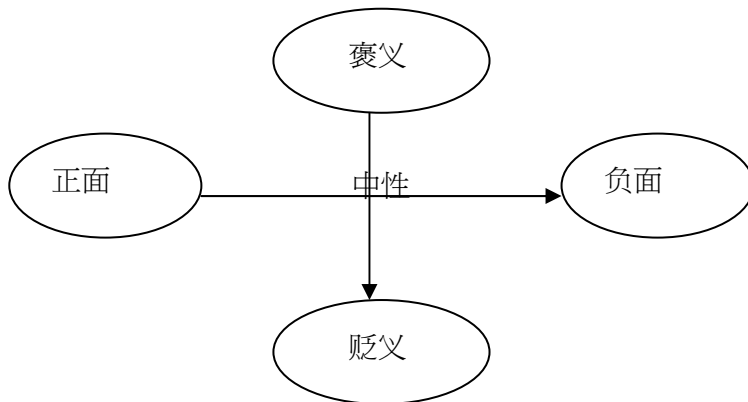
擬聲詞(o)

(1) 嘿嘿 哈哈 喳喳

(2) 喀嚓 嗡嗡

(1)表示“正面或褒義”的詞語，(2)表示“負面或貶義”的詞語。

語法信息詞典的情感屬性設定爲【褒義|正面|中性|負面|貶義】五個級別。



正負面評價是詞彙內容表現的客觀評價，不包含個人的主觀態度。例如：“慘禍”指嚴重的災禍。雖然災禍是指人為不可抗拒的各種災害，具有消極、負面的傾向，但沒有任何人為褒揚或貶損的主觀情緒。正面/負面判斷的方式一方面來自詞義本身，另一方面也來自詞義的上下文環境。詞語的【正面、負面】在表面上不容易判別，需要通過和搭配詞語的共現獲得某種情感傾向，例如“單純”，如果說孩子單純，一般指孩子純潔，表現正面評價；如果說計算機單純，指的就是計算機思想單一，頭腦簡單，四肢發達，表示的是一種負面評價。

【褒義、貶義】和語言學的褒貶基本相當，詞語的褒貶評價從詞彙本身也可以看出，和正負面評價不同的是，【褒義、貶義】的詞語往往帶有人為的主觀因素。例如：“敵人和強盜”就加入了人們對壞人的一種評價態度，因為從詞義本身我們就可以斷定這兩個詞是貶義。除此之外，【褒義、貶義】也傾向於在一組近義詞或反義詞中判定。如“固執和頑固”都表示負面色彩，但兩者有不同程度的差別，“頑固”相對於“固執”貶義的程度更深一些。因此可以設定：固執【+負面】；頑固【+負面 +貶義】。

根據上述標準，我們可以把名詞(1)(2)可以細分為：

【正面】春光

【褒義】慈父 牛市 英雄 勇士 勞模

【負面】癌細胞 愛滋病毒 悲歌 弊端 殘骸 慘禍

【貶義】熊市 黑窩 痞子

詞語的【褒義、貶義】表現出來的情感傾向比較明顯，實際上可以算作【正面、負面】的特殊形式。通常情況下，從詞語的【褒義、貶義】可以推導出詞語的【正面、負面】評價，但是【正面、負面】評價並不一定能推導出詞語的【褒義、貶義】評價。對於【褒義、貶義】這兩種特殊形式，往往並不適合描述上面列舉的全部詞語類別。因此在實際填寫過程中詞類的屬性設定也略有差別。例如：擬聲詞(o)、嘆詞(e)只設定【正面、負面、中性】三個屬性選擇，而且個別還要有所調整，比如代詞(r)“鄙人，吾輩”按照

上述標準不好做出判定，因此評價屬性定義為【敬語|謙語】。

4. 基于語法信息詞典的詞語評價調查

詞語情感評價不僅僅涉及到特定詞或與其搭配的短語本身，還涉及到兩者之間相互作用而產生微妙的情感意義。語義表現出來的正面和負面的感情色彩往往通過直覺發現，但是這種憑直覺發現詞語所具有的感情色彩無法驗證，而基于千萬字量級的漢語基本標注語料庫和語法信息詞典提供的知識足夠使我們把這些現象通過統計揭示出來。例如：語法信息詞典提供的常用詞彙“有”和“味”都是中性詞語，但是它們進入句子框架后卻表現出具有明顯的情感評價傾向。它們在真實語料中有什麼樣的情感表現？它所表現的情感傾向的概率是多少？這些問題我們都可以通過千萬字量級的人民日報語料庫加以驗證。統計發現有以下兩種情況。

第一種情況，“有味”獨立出現，有時后面也加兒化，一共有 12 例

- 1、在街道上唱，公園裏演，蠻有味兒。
- 2、人安靜下來，越看越愛看，越看越有味。
- 3、即使是拈得雞肋者，也會叫道“食之有味，棄之無禮”。
- 4、聽上了廣播，日子是越過越有味了
- 5、做到真實可信，親切有味，引人入勝
- 6、人間有味是清歡——大型畫冊《中國竹工藝》賞析
- 7、可以化腐朽為神奇，變無味為有味。
- 8、先期讀到《雷達散文》的人士，都對雷達的散文評價甚高，認為雷達的散文自然而有味。
- 9、同裏人自己則說，同裏的橋有韻、有味，橋是同裏的性格，
- 10、她們洋溢著自然天成的性情美：自信，進取，活力四射，光彩照人，有為更有味。
- 11、讀書一日，就有一日之益。“讀書有味身忘老。”這是讀書從“苦讀”進到“樂學”境界的表現。
- 12、乘警巡船時在船艙內聞見氣味異常，找到有味車的車主，車主先謊稱是食品添加劑，后又稱是化工品。

第二種情況又包含兩個小類。第一小類“有味”作為固定短語的一部分出現。出現在“津津有味”的例子有 36 例，和“津津有味”共現的動詞一般都是“聽、看、閱讀、喝、講、談起、品嚐、吃、嚼、介紹”等。例如：

- 1、沒想到所有家長都聽得津津有味。
- 2、斯圖拉普一直捧著一本名為《我心依舊》的小說讀得津津有味。
- 3、一邊拿著一本豎排的中文書在津津有味地閱讀。
- 4、正當我津津有味地瞧著時，忽然傳來“當”的一聲，

5、便情不自禁地坐在攤前，匯入津津有味的食客行列。

第二小類，“有味”出現在固定短語“有滋有味”的例子有 28 例。“有滋有味”一般在句子中做補語。與其共現的動詞有“看、喝、吃、啃、說、品味、生活、過、覺、唱、嘗得、干得、打發”。

- 1、別看不是專業演出，可鄉親們依然看得有滋有味。
- 2、她都堅強地挺過來了，而且生活得有滋有味。
- 3、擺碟花生米，就可以有滋有味地喝起來。
- 4、生活也必然會有滋有味，精精神神，充滿無窮樂趣。
- 5、“都都”生活良好，有滋有味地吃了生日蛋糕。

還有一句出現在“有情有味”中。例如：

- 7、他的畫是有情有味有看頭和經得起琢磨的。

以上的所有例句來自于一年半的人民日報語料。“有味”這個短語在出現的所有例句中，76 句有 75 句帶有褒義的感情色彩，約佔所有出現例句的 98.7%。整個句子表現出正面的情感傾向，只有一句是表示負面地含義。即：第一種情況中的例 12 “…有味車…”。“有味車”中的“有味”是指不好的氣味，應該是“有氣味”的簡寫，在所有的語料當中只有一例說的是車上有味的含義，表現出消極的負面評價。而表示正面、褒義情感傾向的“有味”是指“有味道”的含義。我們在語料中找到了 15 個這樣的例子。例如：

- 1、也不論是講述生動有趣的、與飲食有關的故事，總之得有味道，品出多味的生活。
- 2、但該劇的成功，或說最大的特色，即是臺詞很有味道，語言精彩。

它們雖然字形相同，但是所表示的意思是不一樣的。如何在文本中自動判別“有味”屬於哪種類型值得進一步研究。

5. 計算機的形式化研究

語法信息詞典採用成熟的關係資料庫技術，詞語的評價屬性作為原詞典的一個擴充項也沿用此種結構描述。目前語法信息詞典已經把“津津有味”、“有滋有味”收入詞典，而“有味”不作為一個詞條。人民日報基本標注語料庫統一看作一個切分單位。在這方面兩者是不完全等同的。

- 1、越/d 看/v 越/d 愛/v 看/v ，/w 越/d 看/v 越/d 有味/a 。/w
- 2、沒/d 想到/v 所有/b 家長/n 都/d 聽/v 得/u 津津有味/i 。/w
- 3、可/c 鄉親/n 們/k 依然/z 看/v 得/u 有滋有味/l 。

根據這種情況，首先把詞典含有“有味”的四字短語標注上評價屬性。對於“有”和“味味兒”，我們採用分開處理，類似的情況還有很多，例如：“有思想、有氣質、有頭腦、有意見”。不同的名詞和“有”組合所表現的情感傾向是有差別的，因此我們

對“有+N”的形式化描述重點放在后面的N上,根據N和“有”在語料中的真實表現給出他們的情感屬性概率值。

語法信息詞典形式化的描述信息如下：(下表中頻率數值根據一年半人民日報統計得到)

詞語	拼音	詞性	褒貶評價	正負面評價	搭配詞語	概率
津津有味	Jin1jin1you3wei4	i	褒義	正面	聽、看、閱讀、喝、講、談起、品嚐、吃、嚼、介紹	1.0
有滋有味	You3zi1you3wei4	l	褒義	正面	看,喝、吃、啃、說、品味、生活、過,覺,唱、當得、幹得、打發	1.0
有	You3	v	中性	正面	味、味兒	0.917
味	Wei4	n	中性	正面	有	0.917

當然“有味”也可以當作一個詞條收入語法信息詞典,這時對“有味”的描述就和“津津有味、有滋有味”相類似。不過現在將“有味”分開處理,就是要從搭配的角度考察詞語進入句法環境之后所表現出來的情感變化,從方法論的角度為詞語的情感搭配研究提供一種新思路。

6. 詞語情感評價的應用價值

詞語的情感評價研究是機器理解漢語的新拓展,讓機器真正理解自然語言一直是研究者一個遙遠的夢想。很多學者對此都覺得希望渺茫,其中一個十分重要的原因是,機器在自動分析時只是簡單的模式匹配,根本談不上理解,而目前可利用的語義資源也十分有限,詞語的情感評價資源將是一個重要的補充。比如當我們聽到“經濟衰退,股市下滑”的新聞報導時,我們就會由此判斷最近經濟不景氣,而機器無法判斷,但是如果給定“衰退”和“下滑”具有負面情感傾向的屬性特徵,我們就可以利用機器預測股市的發展。詞語的情感評價研究如果能夠利用現有語料庫的豐富資源,給出量化和定性分析,不僅對文本過濾,信息安全、網頁評價的智能化研究有潛在的應用價值,而且對語言學習者也是一個很重要的信息,它能夠幫助學習者在實際語言使用中選擇正確、恰當的詞語。特別是對於母語非漢語的外國留學生而言,如果不理解詞語的情感評價屬性信息,就會在實際交際當中出現錯誤,而這些詞語所表現出來的情感傾向一方面通過實際的交流獲得,另一方面來源于工具書上的信息。但是影響最大的《現代漢語詞典》沒有提供這方

面的屬性。因此語法信息詞典的詞語情感描述研究也會為外國留學生更深刻地理解漢語詞語提供一個重要的語言參考資源。

上面從計算機實現角度對詞語表現出來的情感傾向進行了簡單分類，而且依據大規模語料的證據給予了計量的表示，這詞語的上下文環境對詞語的情感判定起到關鍵作用。因此結合上下文的情感評價研究值得進一步探索。

當然知識庫手工情感標注往往造價很高。專家們試圖尋找各種自動方式的情感分類技術。例如[Pang 2005]提出一種從文本中自動抽取詞語情感多級分類的研究方法。[Takamura *et al.* 2005] 利用電極螺旋模型 (spin model) (相當于詞語情感的正負極) 來抽取詞語的情感極的研究。情感評價分類是自動分析的基礎，在研究初期手工標注工作還是必要的，未來的工作可以考慮利用機器學習等方法實現基于大規模語料的情感詞語的自動抽取。

致謝

筆者在研究過程中，北京大學計算語言所胡景賀同學、譚貽榮同學、呂學強博士、吳云芳博士提出了很好的建議，在此向他們表示衷心的感謝，同時也要特別感謝黃居仁教授提供的重要參考文獻以及對文章的修改建議。

參考文獻

- 大衛·克裏斯特爾[英]，沈家煊譯，《現代語言學詞典》，商務印書館 第4版，2002，pp:290。
- 俞士汶等，《現代漢語語法信息詞典詳解》，清華大學出版社（第2版），2003，pp:40-41。
- 應英等，“漢語情感意義的機器標注研究初探，” *中文信息學報*（第2期），2002。
- 魯紅英，“名人網頁的相關度評價，” *中文信息學報*，2003.5，pp:5。
- 蘇玉梅，“中文網頁褒貶態度的機器評價，” 碩士論文，2004，pp:7。
- 符淮青，《現代漢語詞彙》，北京大學出版社 1985，pp:28。
- 劉叔新，《漢語描述詞彙學》，商務印書館 1990，pp:11。
- 北京大學中文系現代漢語教研室，《現代漢語》，商務印書館，1997，pp:207-208。
- 呂叔湘，《漢語語法分析問題》，商務印書館，1979。
- 呂叔湘，《現代漢語八百詞》，商務印書館，1980。
- Kleinberg, J., and É. Tardos, “Approximation algorithms for classification problems with pairwise relationships: Metric labeling and Markov random fields,” *Journal of the ACM*, 49(5): 2002, pp. 616-639.
- Kushal, D., S. Lawrence, and D.M. Pennock, “Mining the peanut gallery: Opinion extraction and semantic classification of product reviews,” *In Proceedings of WWW*, 2003, pp. 519-528.

- Moshe K., and J. Schler, "The importance of neutral examples for learning sentiment," *In Workshop on the Analysis of Informal and Formal Information Exchange during Negotiations (FINEXIN)*. 2005.
- Palmer, F., ed. "Selected Papers of J.R. Firth," *Landon: Longman*, 1952-1959, 1968.
- Pang, B., and L. Lee, "Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales," *Proceedings of ACL 2005*, pp. 115-124.
- Parington, A., "Patterns and Meanings :Using Corpus for English Language Research and Teaching", 1996, pp. 68
- Sanjiv, D., and M. Chen, "Yahoo! for Amazon: Extracting market sentiment from stock message boards," *In Proceedings of the Asia Pacific Finance Association Annual Conference (APFA)*. 2001.
- Sinclair, J.M., "Looking Up," *London and Glasgow: William Collins*, 1987.
- Sinclair, J.M., "Corpus. Concordance, Collocation," *Oxford University Press*, 1991.
- Takamura, H., T. Inui, and M. Okumura, "Extracting Semantic Orientation of Words Using Spin Model," *Proceedings of ACL 2005*, pp. 133-140.
- T'sou, B. K.T., O.Y. K. wong, W.L. Wong, and T. B.Y. Lai, "Sentiment and Content Analysis of Chinese News Coverage," *In International Journal of Computer Processing of Oriental Languages*, 18.2. 2005, pp. 171-183.

