

# 情绪事件与肇因的语言分析

李逸薇 黄居仁

**提要** 本文以事件结构为出发点研究情绪, 解释文本中情绪与其肇因事件之间的互动。文章指出, 在涉及情绪文本分析的研究中, 情绪和事件之间的关系通常被研究者所忽略, 而事件往往是情绪触发的最重要因素之一。本文基于事件的情绪分析方法, 通过构建情绪肇因事件语料库的方式探索不同情绪与事件之间的关系。研究分析表明情绪类别与肇因事件特征存在着紧密的联系, 而认知标记也可以显性地表达状态变化情绪动词的肇因事件。

**关键词** 情绪 肇因事件 情绪语料库 及物性 认知标记

## A Linguistic Analysis of Emotion and Cause

LEE Sophia Yat Mei and HUANG Chu-Ren

**Abstract** Previous research on emotion analysis often neglects the interactions between emotions and events, regardless of the integral role event plays in emotion studies. In this paper, emotion construction is deemed to be composed of three sequentially structured sub-events, namely cause event (i.e. emotion triggering event), emotion state, and elicited event (i.e. event that is induced by the emotion). We blend the insights from emotion studies in the fields of linguistics, computer science, psychology, etc. and investigate the interactions between emotions and cause events on that basis.

In general, 71% of the instances in the emotion-sentence corpus express emotions, with sentences expressing *happiness* accounting for the most, and *sadness* the least. For each emotion type, approximately 81% of the instances, on average, contain a cause event. It indicates that an emotion frequently co-occurred with its cause event in text. Moreover, it is suggested that cause events tend to be expressed by means of verbal events, and they often occur at the position to the left of the emotion keyword. Therefore, an in-depth analysis of cause event will have great implications for emotion processing in text.

Following Chang et al.'s (2000) dichotomical classification of emotion verbs, emotion verbs are classified into change-of-state emotion verbs and homogeneous state emotion verbs. We investigate the interrelations between emotions and cause events in terms of transitivity and epistemicity. As for the degree of transitivity of cause events, we examine the affectedness of experiencers of the emotions by focusing on three parameters, namely agentivity, kinesis, and participation. The analysis establishes the relationship between emotion types and cause events. The change-of-state and homogeneous emotion verbs share certain cause event features, except for the *fear* emotion. The cause events of *happiness* have the highest degree of transitivity, whereas the cause events of *sadness* have the lowest. The cause events of *anger* and *surprise* share similar features. Unlike other emotions, the cause

event features of *fear* are different for change-of-state and homogenous emotion verbs in terms of the kinesis of events, i.e. motion.

As for epistemicity, we introduce epistemic markers which are verbs that indicate the experiencer's cognitive awareness of the emotion causes. While cause events of homogeneous emotion verbs are rarely found to be introduced by epistemic markers, cause events of change-of-state emotion verbs are often preceded by epistemic markers. It is observed that the cause events of a positive emotion such as *happiness* tend to be more frequently marked by epistemic markers than that of a negative emotion, such as *sadness* or *anger*. It can be explained by the motivation of the experiencer. That is, the higher motivation the experiencer has to express the certainty of the cause events, the more explicit epistemic marking of cause events is allowed.

This paper considers an emotion as a pivot event that links the cause events and the elicited events, and presents the interactions between emotions and cause events using a corpus-based approach. The annotated corpus will provide useful and valuable resources for linguistic analysis as well as natural language processing. Besides, the linguistic account of the emotion-cause interactions will not only contribute to the construction of linguistic model of emotions, but also lay the groundwork for the computational implementations for automatic emotion detection and classification.

**Keywords** emotion, cause event, emotion, transitivity, epistemic marker

## 1. 研究动机: 情感状态与情感事件

语言的功能在于表情( expressive meaning) 与达意( representational meaning)。语言学各领域在达意方面已有很深入的研究。然而, 在表情的功能上, 相关研究相对薄弱。究其原因, 可能是因为近代语言学研究受到结构主义形式二元论( form-meaning dichotomy) 基本假设的主导而忽略了表情与相关的语言学研究。近年来, 语言学( 特别是语义学) 对事件理论的研究建立了达意/知识表达内容与真实世界关系的桥梁, 我们将在这个桥梁上, 发展情绪分析的新架构, 并从语言现象及事件模式出发, 提出一种新的情绪分析方法。

认知心理学或社会学普遍认为情绪通常包括三个要素: 一是人们对外在刺激的观察或想象; 二是由观察到或想象中的刺激而引起的生理与心理状态的改变; 三是针对该刺激所产生的反应行为。以恐惧情绪为例, 恐惧的产生必须先有恐惧对象存在( 不管是真实的还是想象的); 同时, 恐惧的人会有恐惧相关的生理与心理反应; 恐惧产生后, 人们多半会采取远离刺激或寻求保护等自我保护的行为。不难发现, 我们对恐惧或其他情绪的理解建立在对这个熟悉的复杂事件形态的共同经验上。因此, 我们在研究情绪的语言表达方式时, 也需要由这些共同的事件模式出发。此外, 情绪事件分析方法也能更好地解释为什么表情的语言比达意的语言更有效: 达意的语言内容是与行为抽离的, 而表情的语言内容是建立在人类共同经验的反应模式上的。

## 2. 研究背景与文献探讨

### 2.1 情感( sentiment) 与情绪

“情感”与“情绪”这两个近义词( 甚至包括更加口语化的“感情”) 在汉语里通常可以互用, 但最近国际学术界的研究对情感与情绪作了明显而重要的区分。sentiment analysis 着重于情感的极性( polarity) 研究, 也就是集中在人们对某个对象正面或负面的印象或评价上面。计算与认知研究十几年前已经开始进行这个方向的研究, 并且取得了相当不错的成果。emotion analysis 则是针对人的喜怒哀乐等感情的研究, 由于此研究比较复杂, 最近几年才开始有更多的学者关注并投入到此领域。以“情感分析”来翻译 sentiment analysis, 在汉语计算语言学界已有十余年的历史( 许小颖、陶建华 2003; 赵妍妍等 2010; 李寿山、黄居仁 2010) 。鉴于“情感分析”这个词在学术领域已约定俗成, 为了区分 sentiment 与 emotion, 我们按照惯例, 用“情绪”来翻译 emotion( 医疗领域与计算语言学领域共同采用的术语) , 与特指情感极性的“情感”加以区分。

### 2.2 情感与情绪分析

情感分析或意见挖掘( opinion mining) 是目前计算语言学领域广泛关注的研究方向。该方向的研究同本文所研究的情绪分析有一定联系, 但两个研究方向却有着明显的差别。情感分析所关注的是文本所呈现出的对某个对象的个人意见。但个人意见往往具有一定的主观性和复杂性, 因此, 现有的研究多侧重于意见的表达, 并且只分析情感中“正面”或“负面”的态度, 即事物的情感色彩( 如褒义、贬义等) 。具体而言, 情感分析研究主要集中在以下几个方面:

第一, 主观性分类和情感分类: 这两种分类研究是目前情感分析中两个最为广泛的研究课题。主观性分类是指对一篇文本是否含有主观意见进行判断( Wiebe and Wilson 2002; Yu and Hatzivassiloglou 2003; Wilson, et al. 2004; Kim and Hovy 2004; Riloff, et al. 2005) , 而情感分类是指对一篇含有主观意见的文本作出“正面”或“负面”的识别( Pang, et al. 2002; Turney 2002) 。这些研究通常使用基于规则或者监督机器学习的方法完成分类任务。

第二, 情感词典资源建设: 研究人员使用不同的特征和学习算法, 如词的频率、语音卷标( voice label) 、语法或语义等, 结合大规模语料库和机器学习的方法来识别意见单词及其极性。中文研究方面比较重要的有徐琳宏等( 2008) 、李寿山等( 2013) 。

第三, 情感信息抽取: 研究人员以功能/主题为基础识别出同情感表达相关的评价对象或观点持有者( Popescu and Etzioni 2005; Kobayashi, et al. 2007; Titov and McDonald 2008) 。然而, 计算语言学领域在这方面研究的主流一直是以统计方法与机器学习为主, 因而与语言学分析及理论相脱节。当

然,这其中也有少数例外的研究者,例如,Li等(2010a,2010b)分别利用人称/非人称代词的语言分析、语言学极性理论与极性词的概念解决了情感分析词汇统计法的瓶颈问题。本文也将延续这个观点来论证语言学分析与语言学理论在情感/情绪分类中的中心地位。

总体而言,情感分析是目前计算语言学中一个非常活跃的研究领域。除了其重要的理论研究价值外,该研究还和其他技术相结合,具有广泛的实际应用价值,但语言学分析与理论的贡献还有待加强。

与情感分析相比,情绪分析的研究在计算语言学领域还处在初期阶段。情绪分析研究所关注的是文本表达的人类自然情绪(例如:喜欢、生气等)而非简单的正负面情感色彩(例如:褒义、贬义等)。目前,情绪分析的相关研究还相对比较少,具体包括了以下几个重要方面。Hobbs和Gordon(2008)、Mathieu(2005)从词汇语义学角度探讨了情绪分析。Tokuhisa等(2008)基于事件的动机理论使用无监督方法从网络上收集了部分包含情感的日语句子,并在此语料基础上构建了一个自动情感检测和识别系统并应用到对话系统中。遗憾的是,该文收集的情绪语料质量并不高,影响了自动检测和识别的效果。MoodViews系统(Mishne, et al. 2007)在LiveJournal网站收集了大量带心情批注的英语博客语料,并以此作为标注样本训练了一个情绪自动分类系统。Quan和Ren(2009)以中文博客为资源设计了一个情绪类别标注系统,并且标注了基于博客的情绪语料。在这个标注系统中,他们提出了八种不同情绪类别,包括高兴、悲伤、生气等。

然而,情绪分析的相关研究在神经科学和心理学领域早已有了较为深入的探讨。Plutchik(1962)指出尚缺乏一个单一的、集成的、全面的理论来多方位地有效解读何谓情绪。Russell(1980)提出了一种环形的情绪模型。该模型根据一个圆形结构,从愉悦度(valence)和唤起度(arousal)的两维空间概念来分析不同的情绪。Desmet(2002)随后扩展了Russell的情绪模型,提出了与41个产品相关的情绪环形模型。有别于Russell和Desmet的两维空间概念,Plutchik(1994)提出了一个立体的环形模型。该模型定义了情绪强度变化程度并以色轮之间的关系作比拟描述了两两情绪之间的相似程度。具体而言,该模型把八个主要情绪定义为四组对立的概念,即喜悦-悲伤、信任-厌恶、恐惧-愤怒、惊奇-预期。

Turner(1996)进一步把情绪分析为一阶组合和二阶组合,提出了五个基本情绪:乐、怒、哀、惧、惊(在Turner(2000)的情绪模型里,“惊”正式成为基本情绪之一)。他认为两个基本情绪可以结合形成一阶情绪,例如,恐+悲=担心;三种情绪可以组合形成二阶情绪,例如,悲+恐+怒=内疚。该模型和

Plutchik( 1980) 的模型之间的主要区别是情绪的结合涉及主导的基本情绪和少量的另一基本情绪。例如, 耻辱、仇恨和嫉妒的情绪都是愤怒和恐惧的组合, 其差别在于恐惧程度不同。

情绪分析在语言学研究领域通常着重于研究人们如何用语言来表达情绪。情绪词一般分为两种类型, 即表达性情绪和描述性情绪( Kövecses 2000; Potts 2007) 。表达性情绪是一种直接传达情绪的方式, 如“糟糕、哇、呸”等, 而描述性情绪是一种情感关键词描述的方式, 如愤怒、喜悦和悲伤等情绪的表达。例如, “I am so angry, I left my keys in the car!” 属于描述性情绪, 而“Damn, I left my keys in the car!” 则属于表达性情绪。在语音学研究中, 表达性情绪的研究一直是情绪分析的焦点( Abelin and Allwood 2000) 。

### 2.3 事件结构理论

虽然事件结构的研究在哲学和语言学上已经有了一定的发展, 但是在计算语言学领域的研究才刚刚起步。亚里士多德在著作《形而上学》中提到, 每个事件都有它自身的时态和内部的时空结构。Vendler( 1967) 奠定了四种时态动词类( 又名 Aktionsarten) , 包括状态( states) 、活动( activities) 、达成( achievements) 和完成( accomplishments) 。除了 Vendler( 1967) , Dowty( 1979) 、Bach( 1981) 、Moens 和 Steedman( 1988) 与 Jackendoff( 1990) 也分别使用了不同术语来描述事件的各种属性。近些年来, “事件”这个词已被广泛用于包括 Vendler 提出的时态类型中( Tenny and Pustejovsky 2000) 。

在本研究中, 我们所采用的理论框架是生成词库理论( Generative Lexicon Theory, Pustejovsky 1995) 。该理论的前身是 Davidson( 1967) 提出的描绘事件内部结构的理论, 即一个事件包含外部的肇因事件和内部的改变事件。Pustejovsky( 1995) 提出了类似的理论, 认为对事件相关元素的细粒度区别是事件描述的一种有效方式。他提出了一种树状结构, 代表事件的时间顺序和优势约束, 以及其内部包含的事件( 子事件) 。此外, Im 和 Pustejovsky( 2009) 构建了事件结构词汇, 该词汇负责编码 Aktionsart 和子事件的结构框架。他们基于事件的结构定义事件隐含的词汇蕴涵。事件结构包括事件发生前的状态、事件过程和事件发生后的结果状态。这种细粒度事件的描述方式为事件含义的推理和事件文本的识别提供了较大帮助。宋作艳、黄居仁( 2017) 以汉语为研究对象, 运用生成词库理论对汉语的语言现象作了一系列的描述与阐释。

虽然情绪是涉及事件最重要的因素之一, 但已有的研究几乎忽略了情绪和事件之间的关系, 仅有少数研究探讨了人们对事件的反应与情绪之间的关系。特别是在语言学和计算语言学领域, 这方面的研究还没有涉及。

## 2.4 事件与情绪本体(ontology)

在事件本体建设方面, SUMO( Suggested Upper Merged Ontology) 和 DOLCE ( Descriptive Ontology for Linguistic and Cognitive Engineering) 是目前比较好的选择( Borgo, et al. 2010; Niles and Pease 2001), 这两种工具为构建专业本体资源提供了很好的平台。SUMO 是一个基于一阶逻辑的大规模开源本体, 映象了大规模的多语言词典。该本体提供了多种用于新本体开发和应用的实用工具, 同时 SUMO 还同其他重要资源共享了知识, 例如 Wikipedia 和 WordNet。同 SUMO 类似, DOLCE 也是一个基于一阶逻辑的本体, 该本体是一种面向语言及认知工程的描述性本体。除了本体自身提供的基本分类体系外, 它还为本体的扩充和新本体的构建提供了良好的接口。

情绪本体定义情绪结构并且标识情绪之间的互动关系。目前, 一些情绪本体已经构建出来。Obrenovic 等( 2005) 设计了一个以谓语为基础的英语情绪本体。Oltamari( 2006) 采用 Lexipass 方法, 将 WordNet 和 FrameNet 这两个重要的词汇资源结合起来, 构建了一种新的情绪本体。这种一体化资源普遍提高了两个词汇资源的互联互通, 具有良好的用户接口, 从而增强了情绪本体的实用性。此外, 他们还专门为情感和认知领域建立了知识基础的词汇层。López 等( 2008) 融合了网络本体语言( OWL)、DOLCE 和 FrameNet, 开发了一种基于标准本体语言 OWL 的情绪本体。该情绪本体考虑了上下文和多模式联运的元素, 使得该资源能够被轻松共享和扩展。值得一提的是, 在汉语研究方面, Yang 等( 2008) 采用半自动方法, 利用 HowNet 的事件层次构建了汉语情绪本体, 一共包含 5498 个动词概念。

总体而言, 情绪本体的建构一般基于两种工具: WordNet 和 DOLCE。采用 WordNet 的情绪本体把情绪标为额外的语言表达含义, 而采用 DOLCE 的情绪本体却把情绪视为一种完整的认知和体验。知识本体与语言本体研究的交叉, 虽然经黄居仁、李逸薇( 2013) 在《当代语言学》专栏中引进几篇论文, 但是大规模研究仍然有待展开。

## 3. 情绪的事件模式

语言学理论和心理学理论分别为情绪下了相应的定义( Plutchik 1980; Frida 1986; Weiss and Cropanzano 1996), 然而到目前为止, 这些定义并不能引导情绪研究取得明确的共识。尽管如此, 大部分情绪理论都一致认同所有的情绪都是由某个事件激发的。而在语言学关于事件结构的研究中( Pustejovsky 1991), 情绪状态也被视为关键的事件类型。本研究旨在以事件结构为出发点研究情绪, 解释文本中情绪与其肇因事件之间的互动。首先, 我们将根据 Huang( 2010) 提出的观点, 探讨情绪的时间结构, 重点研究情绪和肇因事件之

间的关系，然后对肇因事件进行语义分析。

3.1 基于事件的情绪分析方法

情绪是人类因实际或感知的外部事件激发的普遍认知状态。以往的研究通常把情绪作为一个后状态或是由外部事件触发而产生的单一事件。本文结合事件注释的情感研究，通过集合语言学、计算机科学及心理学等多个领域与情绪相关的研究基础，进一步了解情绪作为事件的概念，进而将取得的成果应用于语言技术层面。在本研究中，我们将情绪的概念同认知经验和潜在后果联系起来，反映出人类行为的强大动机。例如，“恐惧”这种情绪建立在对潜在威胁的认知上，是生存的必要条件。再比如，“欲望”这种情绪在以目的为导向的社会中起着核心作用。总体来说，我们把情绪作为支点，用于连接报道的事件以及经验者可能采取的行动。图 1 给出了情绪和事件的具体关系描述，图 2 以示例说明了图 1 中具体关系的体现。

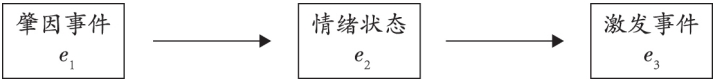


图 1 情绪与事件之间的关系

示例 1

“警方指出，二十三岁的林可晨前天从宜兰北上，原本是要参加校方举办的一个长达七天的活动，但林可晨在参加一天后即脱队，晚间和陈姓同志一同到该 PUB 消费。凌晨约二点半左右，陈姓同志邀约另一名男子在舞池共舞，未料引起林的不悦，两人在 PUB 内吵了起来。据陈姓同志表示，两人争吵后，他气得表明要离开，林可晨即从后追出，双方一路在楼梯间争执，当走到八楼、七楼间，林某突然攀爬上窗台上的栏杆作势往外跳，他上前制止，接着林就掉下，他拉住林的双手，并大呼救命，但终因无力而松手，亲眼目睹林坠楼。”

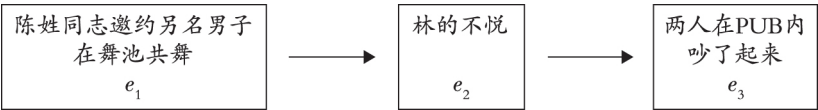


图 2 情绪与事件之间的关系示例

本研究的创新点在于其建立了以情绪作为支点处理事件和情感类别的语言学模型( Lee , et al. 2013; Lee in press) 。一方面，激发情绪的事件可以巩固已提出的情感分类体系的语言学模型。另一方面，情绪描述的分析可以更好地预测事件，从而帮助提升对于文本语境的理解。虽然传统的心理学研究的焦点着重于行为和情绪之间的关系，但是语言学理论并未将两者结合起来进行研究。对于动词形态( 即 Aktionsart ) 和事件结构的研究，语言学家大多集中于探讨事件类型的分类。至于情绪的分析往往着重于词汇语义分解的研究。此外，目前计算语言学广泛关注的情感分析研究，主要集中于评论者对产品的正负面

评价。即使相应的实验结果比较理想，它的应用也局限于某些特定领域，对于更广泛的文本缺乏足够的解释力。

除了情绪的事件结构，本研究同时还关注情绪作为有目的的( telic) 表达以及情绪在文本中的上下文线索。各种情绪都有其目的性的解释，例如，恐惧的出现就是意识到有威胁性的外在条件，因此必须采取行动，以避免受到伤害；幸福的出现就是意识到有理想的外在条件，从而提高生活质量。由此可见，肇因事件与激发事件是了解情绪的最重要因素。不同的情绪(  $e_2$ ) 有不同的肇因事件(  $e_1$ ) 和激发事件(  $e_3$ ) 。这些相关事件一般会出现在情绪事件的上下文中，具体实例可以参考图 2。

在理论分析的基础上，本研究还将人工标注包括情绪和事件信息的大规模语料。

### 3.2 情绪语料库

情绪分析研究往往需要大规模的情绪语料库作支撑，但遗憾的是，迄今为止很少有学者去标注大规模的情绪语料库。其原因在于标注大规模情绪语料库需要耗费大量的时间和精力，同时还涉及语料库资源的选择、收集与整理以及标注体系的制订及标注。与此同时，情绪标注往往因人而异，具有很大的争议性，从而导致语料库的质量难以保证。情绪语料库建设的另外一个困难在于无情绪句子的收集，这方面的句子对于情绪的识别和分类具有关键作用。本研究克服了这些困难，使用一种非监督方法来创建大规模汉语情绪语料库和中性句子语料库，以满足情绪计算的需要( Chen, et al. 2009) 。我们采用了两个不同的语料库去标注有情绪句子，并给出了一些情绪分布的定量分析。

情绪文本的分布同语料库的文体类型密切相关。例如，偏文学的语料往往包含较多的情绪信息。为了避免某些特定文体类型对情绪分析的影响，本文首先使用“中研院”平衡语料库( Chen, et al. 1996) 作为基础语料库，这个语料库的总字数是 1000 万字。此外，我们还使用了一个带词性标注的大规模新闻语料库，即 Tagged Chinese Gigaword Corpus ( Graff and Chen 2003; Huang 2009) 。该语料库有 14 亿个词，文本分别是来自内地( 新华社)、台湾地区( “中央社”) 和新加坡《联合早报》三地收集而来，充分考虑到了汉语的地域差别。在标注语料之前，我们对这三个语料库分别做了自动分词和词性标注的工作。我们用来分析的语料在语法及词汇上可能有区域性的差异，但对本研究而言，更重要的是这些语料中情绪的语言表达与分类是否有一致性。我们的研究分析证实了它们具有一致性。

情绪语料库的具体创建方式如下。首先，我们对每一个情绪关键词用正则表达式去提取语料中含有该情绪关键词的句子，在抽取过程中，要求这个情绪

词必须是动词。然后，用正则表达式去发现这个情绪词前面的词是否为否定词（如“不”）。如果是，则把该句剔除，这主要是因为带否定的情绪关键词表达的情绪类别歧义性较大，加入这些句子会影响我们收集的语料的质量。最后，每个句子的情绪类别可以根据该句子含有的情绪关键词的情绪类别决定。在做情绪自动分析的时候有一点需要注意，即句子中的情绪关键词会被删除掉，这样情绪语料库只能为情绪自动分析提供相应的语境。

上文还提到情绪自动识别还需要无情绪句子语料库，即中性句子语料库。Tokuhisa 等(2008)提出，对情绪计算来说，中性句子语料库的创建是非常重要的，但中性句子的收集并非易事。例如，在日语的网络语料中大约有 80% 的句子是有情绪的。为了了解中性句子在汉语书面表达中的分布，我们采用一个比较直观的方法来自动提取中性句子。给定一个句子，如果它的前后句都不含任何一个情绪词，那么该句子被归为中性。最后，情绪语料库和中性句子语料库共同组成整个以情绪计算为目的的语料库。

#### 4. 情绪肇因事件的语法语意证据

##### 4.1 情绪肇因事件

在有情绪句子中，情绪状态与导致其情绪发生的可能原因建立起了联系。在本研究中，与情绪状态有密切关系的事件被称为肇因事件( Lee, et al. 2009)。肇因事件并不一定实际触发情绪或导致某种情绪的出现。相反，它可以是实际触发情绪的事件或对触发情绪事件的感知。本文的肇因事件是指与情绪有密切关系事件的具体语言表达。我们把肇因事件归纳为两类：动词事件和名词事件。其中动词事件是指事件涉及动词(即命题)，如(1)，或名词化，如(2)，而名词事件仅仅是名词，如(3)。

- (1) 没想到她说的都是真话，让他震惊不已。
- (2) 他对这个充满浓厚爱意的想法高兴得手舞足蹈。
- (3) 尼奥的话很令凯洛琳伤心。

在语料库中我们发现在许多情况下，肇因事件存在指代现象。如在(4)中，愤怒的肇因事件“此”指代的是其先行词“台湾当局以政治理由拒绝当事人返台”。在这种情况下，两者都标记为肇因事件，前者为名词肇因事件，后者为动词肇因事件。假如先行词不出现在上下文内，如(5)，那么就只标注出现的肇因事件“此喜讯”。

- (4) 台湾当局以政治理由拒绝当事人返台。当事人对此愤懑。
- (5) 夫妻俩昨天都对此喜讯兴奋不已。

在实际分析的过程中，真正的肇因事件有时非常复杂，可能同时涉及到多个事件，为了简单起见，我们仅把最接近情绪词的相关事件标注为肇因事件。

4.2 情绪肇因事件语料库

根据 Lee 等( 2012) 制定的情绪关键词列表, 我们从“中研院”平衡语料库中提取出了 6058 个句子, 然后对每个句子进行标注。每一个句子的标注包含三个部分, 第一部分是含有情绪关键词的重点句, 标记为“<FocusSentence>”; 第二部分是重点句的前一句, 标记为“<PrefixSentence>”; 第三部分是重点句的后一句, 标记为“<SuffixSentence>”。情绪关键词分别来自五个基本情绪类别, 考虑到每个情绪的句子分布严重不平衡( 例如, “快乐”的关键词比其他情绪类别多很多), 我们设置了 1600 句的上限。

在标注的过程中我们注意到, 情绪关键词的出现并不意味着一定存在情绪表达。为了确保语料库的质量, 我们剔除了不涉及情绪的句子, 这些句子包括否定句、条件句以及情绪关键词歧义性较大的句子, 最终语料库共获取了 5964 个句子。

为了确保语料库的准确性, 标注工作由两位标注者同时对每个有情绪句子人工标注情绪关键词的肇因事件。由于一个有情绪句子可能含有多于一个的情绪, 情绪关键词将按照<emotionword id=0><emotionword id=1>标记, 依此类推。

573 Y 0/伤心/Sadness  
<PrefixSentence> 妈妈 也 问 了 邻居, 但是 没有 人 见到 小白。  
</PrefixSentence> <FocusSentence>为了 [\* 01n] 这件事 [\* 02n], 我 曾  
<emotionword id=0>伤心</emotionword> 了 好久, 但 也 无济于事。  
</FocusSentence> <SuffixSentence>每 当 再 看到 白色 的 弃 狗, 不 禁  
会 想起 小白。 </SuffixSentence>

图 3 情绪肇因事件的标注实例

图 3 是情绪语料库中的一个标注例子, 其中情绪关键词标为<id=0>, 肇因事件的开始标为[\* 01n], 而肇因事件的结束标为[\* 02n]。n 表示肇因事件为名词事件。情绪关键词有时可以与不止一个肇因事件相关, 在这种情况下, 所有出现在句子中的肇因事件都需要标注。如果有情绪句子没有明确表示的肇因事件, 则不需要做任何标注。

在肇因事件语料库中, 71%的句子表达情绪。表达“快乐”情绪的句子最多, 共 1327 句。表达“悲伤”情绪的句子最少, 共 616 句。对于每一个情绪类别, 平均约 81%的有情绪句子包含一个肇因事件。其中, 表达“快乐”和“惊喜”的有情绪句子中肇因事件比重最高( 约 85%), 而在表达“愤怒”情绪的句子中, 肇因事件比重最低( 约 74%)。此外, 通过分析语料发现, 肇因事件倾向于以动词事件来表达, 并且往往出现在情绪关键词的左侧( 详细讨论可参阅 Lee 2012, 2013)。可见, 肇因事件通常会与情绪事件同时出现在上下文。因此, 对肇因事件进行深入的研究有助于文本情绪表达的分析。

4.3 情绪肇因事件的语法语义分析

通过上文的研究和分析，我们可以认为情绪是一种特殊事件类型，它是由实际的或者可预见的事件(即肇因事件)触发的。为了构建一个完善的情绪事件模型，我们做了一系列相关研究，并基于此深入探讨了面对不同情绪的肇因事件的语法和语义特性。在这些研究工作中，我们着重关注五个基本情绪(即乐、悲、悔、恐、惊)及其与肇因事件之间的相互关系，并重点考察肇因事件的两个特征：及物性和认知标记( Lee 2010 , in press)。

4.3.1 肇因事件的及物性

为了深入分析情绪及其与肇因事件的关系，我们从语义角度观察肇因事件的及物性是否与情绪类别有紧密关系。按照 Hopper 和 Thompson( 1980) 的及物性定义，我们特别关注及物性的三个参数，即施事性、动态性和参与性。依据这些参数，考察了情感体验者如何受肇因事件的影响。三个参数的简要定义如下：

施事性：根据 Dowty( 1991) 对角色的定义，参与角色可以分为原型施事者( Proto-Agent) 和原型受事者( Proto-Patient) 两种。肇因事件主语的施事性表现了其影响情绪经验者的意图。当肇因事件的主语倾向于施事者属性时，显示其有较强的意图去影响情绪体验者。

动态性：肇因事件可以分为两种类型，即行动型和非行动型( Talmy 2000) 。行动型肇因事件往往直接影响情绪经验者，而非行动肇因事件则往往间接影响情绪经验者。

参与性：考察情绪经验者是否参与肇因事件，参与性体现了经验者受影响的程度。参与肇因事件的经验者偏向于直接受影响，没有参与的经验者则偏向于间接受影响。

针对以上三个参数，我们分析了 1000 句带有肇因事件的有情绪句子，考察了这些句子中各种情绪在三个参数上面的表现。分析结果显示情绪类别和肇因事件之间存在着一定的关系。表 1 所示的是各种情绪类别的肇因事件特征，除了“恐”这种情绪之外，其他情绪的状态变化动词与非状态变化动词都有相同倾向的肇因事件特征。

情绪类别	肇因事件特征
乐	[+原施事] [+动态性] [+参与性]
悲	[+原施事] [-动态性] [-参与性]
恼	[+原施事] [+动态性] [-参与性]
恐	状态变化动词 [+原施事] [-动态性] [+参与性]
	非状态变化动词 [+原施事] [+动态性] [+参与性]
惊	[+原施事] [+动态性] [-参与性]

表 1 情绪类别的肇因事件特征

通过对表 1 的分析可以看出在五种基本情绪中,“乐”的肇因事件具有最强的及物性。换言之,“乐”的肇因事件往往会包括施事主语、动态和经验者。与之相反的是“悲”,其肇因事件较为不及物。“恼”和“惊”有着类似的肇因事件特征:有原施事主语、动态性和经验者不参与性。相对而言,“恐”有别于其他情绪类别。状态变化动词与非状态变化动词有着不一样的肇因事件特征,两者的差别在于事件的动态性。我们发现,状态变化动词往往跟潜在的肇因事件相关,而非状态变化动词则与实际的肇因事件相关。潜在的肇因事件通常是经验者想象出来的事件,可能会发生,也可能不会发生。例如,害怕孤独、害怕老虎等各种原因,这些通常不涉及动态性。实际肇因事件大多涉及动态性,如被老虎袭击、被绑架等。

我们通过考察肇因事件和情绪类别之间的相互作用,详尽分析了各种不同情绪类别的肇因事件的语义特征,这些特征可以应用于情感计算中,例如:情绪识别和情绪分类等任务。

#### 4.3.2 情绪肇因事件的认知标记

从语法的角度,我们又考察了识别情绪动词的肇因事件的认知手段。Chang 等(2000)指出状态变化情绪动词一般直接以肇因事件作为补语,如(6)所示。

(6) 我很高兴他们回来了。

通过分析语料,我们也发现很多状态变化情绪动词的肇因事件都由一些认知动词引出,如(7)中的“看见”。因此,我们称这些动词为认知标记。

(7) 我很高兴看见他们回来了。

在认知标记方面,本研究主要探讨了两个重要问题:(1)我们到底是如何通过认知标记识别状态变化情绪动词的肇因事件?(2)是什么原因推动了这样一种认知标记的使用?

针对这两个问题,我们重点研究了五种基本情绪的语料。根据 Chang 等(2000)的识别状态变化情绪动词的测验,我们得出了五对相对应的有代表性的汉语词汇:高兴/快乐、伤心/悲伤、害怕/恐惧、生气/愤怒、惊讶/震惊。这些词汇都是在语料中出现次数最多的词语。我们的数据是从 Chinese Gigaword Corpus(中文十亿词语料库)中抽取的,该语料库是目前规模最大的中文语料库,有 14 亿中文语料。然后利用 Chinese Word Sketch(Huang, et al. 2005)工具从语料中自动生成语言学上有意义的搭配,该工具提供了丰富的基于词汇的各种语法关系并且带有相应的统计信息。

我们归纳了五种认知标记(见表 2):视觉性、听觉性、认知性、发现性和存在性。表 2 给出了具体类别的认知标记及相应的例句。在标注语料的基础

上，我们统计了各种认知标记类别在各个状态变化动词中出现的频率，统计结果如表 3 所示。从表 3 中可以发现在状态变化动词中，“高兴”带认知标记的频次最高，尤其是视觉性的认知标记。“伤心”和“生气”带认知标记的频次最低。非状态变化动词，如“快乐”和“悲伤”，基本上不带任何含认知标记的补语。

类别	认知标记	例句
视觉性	看、见、看见、看到、见到	他很高兴看到中国大陆政局稳定。
听觉性	听、听到、听说	我很高兴听到很多关于警察行动的反应受到人民的注意。
认知性	知道、得知、得悉、获知、获悉	他很高兴得知香港与台湾间的旅游能有进展。
发现性	发现	她很高兴发现女性歌迷比以前增加很多。
存在性	有	他很高兴有很多对话已出现并落实。

表 2 认知标记的类别与相应实例

认知标记	高兴	伤心	害怕	生气	惊讶
视觉性	2462	2	8	0	21
听觉性	123	0	4	0	2
认知性	137	2	0	0	3
发现性	16	0	1	2	219
存在性	755	1	43	1	17
总认知标记	3493	5	56	3	262
总数	41043	5967	7989	3315	6892
归一化比率	8.50%	0.07%	0.70%	0.09%	3.80%

表 3 状态变化动词的认知标记统计

通过以上分析我们认为，认知标记的作用是表示情绪肇因事件的认知意识。“高兴”的肇因事件可以由五种认知标记引出的补语显性表达出来。统计结果表明，“生气”和“伤心”的语句表达中极少出现认知标记，这种现象可以用波丽安娜效应( Pollyanna Effect) 或积极偏向来解释。

按照情绪所表达的情感极性，情绪一般可以分为愉悦情绪和不良情绪两种。虽然这种分类跟文化背景有很大的关系，但是大部分情绪在不同的文化背景下仍然有相同的极性分类。例如，“乐”这种情绪基本被认为是正面的、令人愉悦的，而“悲”则是一种负面的不良情绪。对于像“乐”之类的愉悦情绪，人们在文本表达时倾向于使用认知标记，因为情绪体验者往往有强烈的动机去识别引发这种情绪的事件，而这些事件通常是可以感知的。相对来说，像“悲”之类的不良情绪允许的认知标记则十分有限，其原因在于这种情绪往往来自于体验者内在的、自发的动因，不需要为肇因事件提供明确的证据。综上所述，状态变化情绪动词的肇因事件可由认知标记引出的补语显性地表达出来，正面情绪往往利用认知标记明确地标记肇因事件，负面的情绪则相反。

## 5. 结论

本研究以情绪为焦点,将与情绪密切联系的肇因事件和激发事件作为情绪的两种特殊角色进行研究,旨在为情绪分析研究提出一个以语言学本体为出发点的方法,并证明本体研究能协助计算机处理应用研究。本研究的理论核心是情绪的语言事件模型,通过构建中英文的情绪分类体系以及相应的情绪语料库建立起一个完善的情绪事件模型,这些语料将为未来的语言分析以及计算语言学分析提供宝贵的资源。此外,本文利用该情绪语料库对情绪的肇因事件进行了深入的语言学研究,分析结果不仅对建立情绪的语言学模式有极大的帮助,同时也为情绪自动识别和分类系统提供了有效依据。

## References [引用文献]

- Abelin, Åsa and Jens Allwood. 2000. Cross-linguistic interpretation of emotional prosody. In *Proceedings of the ISCA ITRW on Speech and Emotion*. Pp.110-3.
- Bach, Emmon W. 1981. On time, tense, and aspect: An essay in English metaphysics. In Peter Cole, ed., *Radical Pragmatic*. New York: Academic Press. Pp.62-81.
- Borgo, Stefano, Massimiliano Carrara, Pawel Garbacz, and Pieter E. Vermaas. 2010. Formalizations of functions within the DOLCE ontology. In *Proceedings of the Eighth International Symposium on Tools and Methods of Competitive Engineering (TMCE 2010)*, Vol. 1. Pp.113-26.
- Chang, Li-li (张丽丽), Keh-Jiann Chen (陈克健), and Chu-Ren Huang (黄居仁). 2000. Alternation across semantic field: A study of Mandarin verbs of emotion. *Computational Linguistics and Chinese Language Processing* 5, 1: 61-80.
- Chen, Keh-Jiann, Chu-Ren Huang, Li-ping Chang (张莉萍) and Hui-Li Hsu (许蕙丽). 1996. Sinica corpus: Design methodology for balanced corpora. In *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation*. Pp.167-76.
- Chen, Ying (陈瑛) and Sophia Y.-M. Lee (李逸薇) and Chu-Ren Huang. 2009. A cognitive-based annotation system for emotion computing. *Proceedings of the Third Linguistic Annotation Workshop (The LAW III)*. Pp.1-9.
- Davidson, Donald. 1967. The logical form of action sentences. In *Essays on Actions and Events*. Oxford: Clarendon Press. Pp.105-48.
- Desmet, Pieter. 2002. Designing emotion. Ph. D. diss., Delft University of Technology, Delft, the Netherlands.
- Dowty, David. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Springer Netherlands.
- . 1991. Thematic proto-roles and argument selection. *Language* 67, 3: 547-619.
- Frijda, Nico H. 1986. *The Emotions: Studies in Emotion and Social Interaction*. Paris: Maison de Sciences de l'Homme; Cambridge University Press.
- Graff, David and Ke Chen. 2003. Chinese Gigaword Corpus. Linguistic data consortium, Philadelphia.
- Hobbs, Jerry R. and Andrew S. Gordon. 2008. The deep lexical semantics of emotions. The Sixth International Conference on Language Resources and Evaluation (LREC) Workshop on Sentiment Analysis: Emotion, Metaphor, Ontology and Terminology (EMOT), Marrakech, Morocco.
- Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, 2: 251-99.
- Huang, Chu-Ren. 2009. Tagged Chinese Gigaword Version 2. 0. Linguistics data consortium,

Philadelphia.

- . 2010. Towards an event-based theory of emotion classification. The 6th Workshop on Ontologies and Lexical Resources ,Beijing ,China.
- Huang ,Chu-Ren ,Adam Kilgarrriff ,Yiching Wu ,Chih-Ming Chiu ,Simon Smith ,Pavel Rychly ,et al. 2005. Chinese sketch engine and the extraction of grammatical collocations. *Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing*. Pp.48-55.
- Huang ,Chu-Ren and Sophia Y.-M. Lee. 2013. Knowledge systems and the construction of knowledge systems: Introducing interdisciplinary studies on ontologies and language sciences. *Contemporary Linguistics* 3: 284-93. [2013 ,知识的系统与知识系统的建构: 知识本体语言科学整合研究 《当代语言学》第3期 ,284-93 页。]
- Im ,Seohyum and James Pustejovsky. 2009. Annotating event implicatures for textual inference tasks. The 5th International Conference on Generative Approaches to the Lexicon ,Pisa ,Italy.
- Jackendoff ,Ray. 1990. *Semantic Structures*. Cambridge ,MA: The MIT Press.
- Kim ,Soo-Min and Eduard Hovy. 2004. Determining the sentiment of opinions. In *Proceedings of the International Conference on Computational Linguistics ( COLING 2004)* . Pp.1367-73.
- Kobayashi ,Nozomi ,Kentaro Inui ,and Yuji Matsumoto. 2007. Extracting aspect-evaluation and aspect-of relations in opinion mining. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning ( EMNLP-CoNLL 2007)* . Pp.1065-74.
- Kövecses ,Zoltán. 2000. *Metaphor and Emotion: Language ,Culture and Body in Human Feeling*. Cambridge: Cambridge University Press.
- Lee ,Sophia Y.-M. 2010. A linguistic approach to emotion detection and classification. Ph.D. diss. , The Hong Kong Polytechnic University ,Hong Kong.
- . In press. *Emotion and Cause: Linguistic Theory and Computational Implementation*. Berlin: Springer.
- Lee ,Sophia Y.-M ,Ying Chen ,and Chu-Ren Huang. 2009. Cause event representations for happiness and surprise. *Proceedings of the 23rd Pacific Asia Conference on Language ,Information and Computation*. Pp.297-306.
- Lee ,Sophia Y.-M. ,Huarui Zhang ( 张化瑞) ,and Chu-Ren Huang. 2013. An event-based emotion corpus. In Pengyuan Liu ( 刘鹏远) and Qi Su ( 苏祺) ,eds. ,*Proceedings of the 14th Chinese Lexical Semantics Workshop*. Berlin / Heidelberg: Springer. Pp.635-44.
- Lee ,Sophia Y.-M. ,Ying Chen ,Chu-Ren Huang ,and Shoushan Li ( 李寿山) . 2012. Detecting emotion causes with a linguistic rule-based approach. *Computational Intelligence ( Special Issues on Computational Approaches to Analysis of Emotion in Text)* 29 3: 390-416.
- Li ,Shoushan and Chu-Ren Huang. 2010. Chinese sentiment classification based on stacking combination method. *Journal of Chinese Information Processing* 24 ,5: 56-61. [2010 ,基于 Stacking 组合分类方法的中文情感分类研究 《中文信息学报》第5期 ,56-61 页。]
- Li ,Shoushan ,Chu-Ren Huang ,Guodong Zhou ( 周国栋) ,and Sophia Y.-M. Lee. 2010a. Employing personal/impersonal views in supervised and semi-supervised sentiment classification. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics ( ACL 2010)* . Pp.414-23.
- Li ,Shoushan ,Sophia Y.-M. Lee ,Ying Chen ,Chu-Ren Huang ,and Guodong Zhou. 2010b. Sentiment classification and polarity shifting. In *Proceedings of the 23rd International Conference on Computational Linguistics ( COLING 2010)* . Pp.635-43.
- Li ,Shoushan ,Sophia Y.-M. Lee ,Chu-Ren Huang ,and Yan Su ( 苏艳) . 2013. Construction of Chinese sentiment lexicon using bilingual information and label propagation algorithm. *Journal of Chinese Information Processing* 6: 75-82. [2013 ,基于双语信息和标签传播算法的中文情感词典构建方法 《中文信息学报》第6期 ,75-82 页。]
- López ,Juan M. , Rosa Gil ,Roberto García ,Idoia Cearreta ,and Nestor Garay. 2008. Towards an ontology for describing emotions. In Miltiadis D. Lytras ,John M. Carroll ,Ernesto Damiani ,and

- Robert D. Tennyson , eds. , *Emerging Technologies and Information System for the Knowledge Society*. Berlin / Heidelberg: Springer. Pp.96–104.
- Mathieu , Yvette Yannick. 2005. Annotations of emotions and feelings in texts. In Jianhua Tao ( 陶建华 ) , Tieniu Tan and Rosalind W. Picard , eds. , *Proceedings of Conference on Affective Computing and intelligent Interaction ( AII 2005 )* . Berlin / Heidelberg: Springer. Pp.350–7.
- Mishne , Gilad , Krisztian Balog , Maarten Rijke , and Breyten Ernsting. 2007. Mood Views: Tracking and searching mood-annotated blog posts. In *Proceedings of International Conference on Weblogs and Social Media ( ICWSM-2007 )* . Pp.323–4.
- Moens , Marc. and Mark Steedman. 1988. Temporal ontology and temporal reference. *Computational Linguistics* 14 2: 15–28.
- Niles , Ian and Adam Pease. 2001. Towards a standard upper ontology. *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems ( FOIS-2001 )* . Pp.2–9.
- Obrenovic , Zeljko , Nestor Garay , Juan Miguel López , Inmaculada Fajardo , and Idoia Cearreta. 2005. An ontology for description of emotional cues. In *Proceedings of Affective Computing and Intelligent Interaction ( AII 2005 )* . Pp.505–12.
- Oltamari , Alessandro. 2006. LexiPass methodology: A conceptual path from frames to senses and back. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation ( LREC 2006 )* . Pp.2311–4.
- Pang , Bo , Lilian Lee and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing ( EMNLP )* . Pp.79–86.
- Plutchik , Robert. 1962. *The Emotions: Facts , Theories and a New Model*. New York: Random House.
- . 1980. *Emotions: A Psychoevolutionary Synthesis*. New York: Harper & Row.
- . 1994. *The Psychology and Biology of Emotion*. New York , NY: Harper-Collins.
- Popescu , Ana-Maris and Oren Etzioni. 2005. Extracting product features and opinions from reviews. In *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing ( HLT/EMNLP )* . Pp.339–46.
- Potts , Christopher. 2007. Expressive dimension. *Theoretical Linguistics* 33 2: 165–98.
- Pustejovsky , James. 1991. The syntax of event structure. *Cognition* 41 1: 47–81.
- . 1995. *The Generative Lexicon*. Cambridge , MA: The MIT Press.
- Quan , Changqin and Fuji Ren ( 任福继 ) . 2009. Construction of a blog emotion corpus for Chinese expression analysis. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing ( EMNLP )* . Pp.1446–54.
- Riloff , Ellen , Janyce Wiebe , and William Phillips. 2005. Exploiting subjectivity classification to improve information extraction. *Proceedings of the 20th National Conference on Artificial Intelligence ( AAAI-05 )* . Pp.1106–11.
- Russell , James A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39 6: 1161–78.
- Song , Zuoyan ( 宋作艳 ) and Chu-Ren Huang. 2017. *Generative Lexicon Studies in Chinese*. Beijing: The Commercial Press. [2017, 《生成辞汇理论与汉语研究》。北京: 商务印书馆。]
- Talmy , Leonard. 2000. *Toward a Cognitive Semantics* , Vols. 1 and 2. Cambridge , MA: The MIT Press.
- Tenny , Carol and James Pustejovsky , eds. 2000. *Events as Grammatical Objects*. CSLI Publications.
- Titov , Ivan and Ryan McDonald. 2008. A joint model of text and aspect ratings for sentiment summarization. In *Proceedings of 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies ( ACL-08: HLT )* . Pp.308–16.
- Tokuhisa , Ryoko , Kentaro Inui , and Yuji Matsumoto. 2008. Emotion classification using massive examples extracted from the web. *Proceedings of the 22nd International Conference on Computational linguistics ( COLING 2008 )* . Pp.881–8.
- Turner , Jonathan H. 1996. The evolution of emotions in humans: Darwinian-Durkheimian analysis. *Journal for the Theory of Social Behaviour* 26 1: 1–33.

- . 2000. *On the Origins of Human Emotions: A Sociological Inquiry into the Evolution of Human Affect*. Stanford, CA: Stanford University Press.
- Turney, Peter D. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL)*. Pp.417–24.
- Vendler, Zeno. 1967. Verbs and times. In *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press. Pp.97–121.
- Weiss, Howard M. and Russell Cropanzano. 1996. An affective events approach to job satisfaction. In Barry M. Staw and Larry L. Cummings, eds., *Research in Organizational Behavior*. Greenwich, CT: JAI Press. Pp.1–74.
- Wiebe, Janyce and Theresa Wilson. 2002. Learning to disambiguate potentially subjective expressions. In *Proceedings of the 6th Conference on Natural Language Learning*. Pp.1–7.
- Wilson, Theresa, Janyce Wiebe, and Rebecca Hwa. 2004. Just how mad are you? Finding strong and weak opinion clauses. In *Proceedings of the 19th National Conference on Artificial Intelligence and Sixteenth Conference on Innovation Applications of Artificial Intelligence*. Pp.761–7.
- Xu, Linhong (徐琳宏), Hongfei Lin (林鸿飞), Yu Pan (潘宇), Hui Ren (任惠), and Jianmei Chen (陈建美). 2008. Constructing the affective lexicon ontology. *Journal of the China Society for Scientific and Technical Information* 2: 180–5. [2008, 情感词汇本体的构造. 《情报学报》第2期, 180–5页。]
- Xu, Xiaoying (徐小颖) and Jianhua Tao. 2003. The study of affective categorization in Chinese. In *Proceedings of 1st Chinese Conference on Affective Computing and Intelligent Interaction*. Pp.199–205. [2003, 汉语情感系统中情感划分的研究. 见《第一届中国情感计算及智能交互学术会议论文集》。199–205页。]
- Yang, Jiajun, David B. Bracewell, Fuji Ren, and Shingo Kuroiwa. 2008. The creation of a Chinese emotion ontology based on HowNet. *Engineering Letters* 16, 1: 166–71.
- Yu, Hong and Vasileios Hatzivassiloglou. 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 (EMNLP)*. Pp.129–36.
- Zhao, Yanyan (赵妍妍), Bin Qin (秦兵), and Ting Liu (刘挺). 2010. Sentiment analysis. *Journal of Software* 8: 1834–48. [2010, 文本情感分析. 《软件学报》第8期, 1834–48页。]

#### 第一作者简介

李逸薇, 女, 博士, 香港理工大学中文及双语学系助理教授。研究兴趣: 中国语言学、语料库语言学、计算语言学。代表作: *Emotion and Cause: Linguistic Theory and Computational Implementation*。电子邮件: ymlee@polyu.edu.hk

LEE Sophia Yat Mei, female, Ph.D., is an assistant professor at the Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University. Her research interest includes Chinese linguistics, corpus linguistics, and computational linguistics. Her major publication is: *Emotion and Cause: Linguistic Theory and Computational Implementation*. E-mail: ymlee@polyu.edu.hk

#### 作者单位及通信地址

香港理工大学中文及双语学系

#### Affiliation and Contact

Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong