

基于语料库的新媒体语言透视

崔希亮

(北京语言大学汉语国际教育研究院,北京 100083)

提 要 新媒体语言有自己的特点,很多人都做过研究。但是基于语料库的研究还不多见。本文以 BCC 语料库为根据,统计考察新媒体各种语言现象,并透视它们的语言学价值、修辞学价值以及社会学价值。

关键词 新媒体语言 基于语料库 语言学 修辞学 社会学

DOI:10.16027/j.cnki.cn31-2043/h.2019.05.001

所谓新媒体指的是网络媒体,包括电脑网络媒体和移动网络媒体,与传统的平面媒体(报纸杂志)和有声媒体(广播电视)相对而言。新媒体是一种新的传播媒介,传播速度快,传播内容庞杂,传播范围广。新媒体语言最主要的特点是交互性、即时性和流行性,还有随意性和私密性,也有一些研究者提出了其他一些特点,如时尚、新潮、创新性、个性化、简洁、粗俗等等,不一而足(李琼 2016: 23-24; 罗譞 2018: 59-61; 霍慧娜 2017: 4-5)。有关新媒体语言的研究不少,但是仍有深入研究的空间。有人对 2002 到 2012 年新媒体的语言研究做了总结回顾(赵雪、曹彦男 2012: 81-85),可见学术界早已关注新媒体的语言问题。2012 年以后不断有新的研究成果问世(尚伟 2017: 11-15),尤其是有学者从话语应用和传播的角度对特定领域的新媒体语言进行语用分析,有一些新的发现(王建华 2018: 101-108),王建华(2017)有专书出版,还有人从社会环境和语言选择的角度对特定领域新媒体语言进行研究(施麟麒 2018: 120-127)。大家角度不同,研究方法不同,所观察到的现象也有所不同。有学者从传播方式的角度出发,认为人类社会经历了几次语言革命,中国的语言自 1919 年已经发生了两次语言革命,新媒体的普遍应用是第三次语言革命(赵莉 2014: 2-3)。(见图 1)

李雪(2016: 9)有一个调查,认为新媒体有如下一些特点(见图 2)。

尽管新媒体语言的影响力很大,但是它们基本上局限在特定的场合、特定的人群(朋友圈、粉丝圈、网游圈、韩流圈等),说媒体语言革命也许为时尚早。我们以北京语言大学的 BCC 语料库^①为依据,比较一些新媒体语言现象在不同语料库中的分布,从而透视新媒体语言在语言学、修辞学和社会学中的价值。

一、新媒体语言的语言学价值

新媒体语言的产生和传播与新媒体的传播方式有关。新媒体的传播与平面媒体不同:传

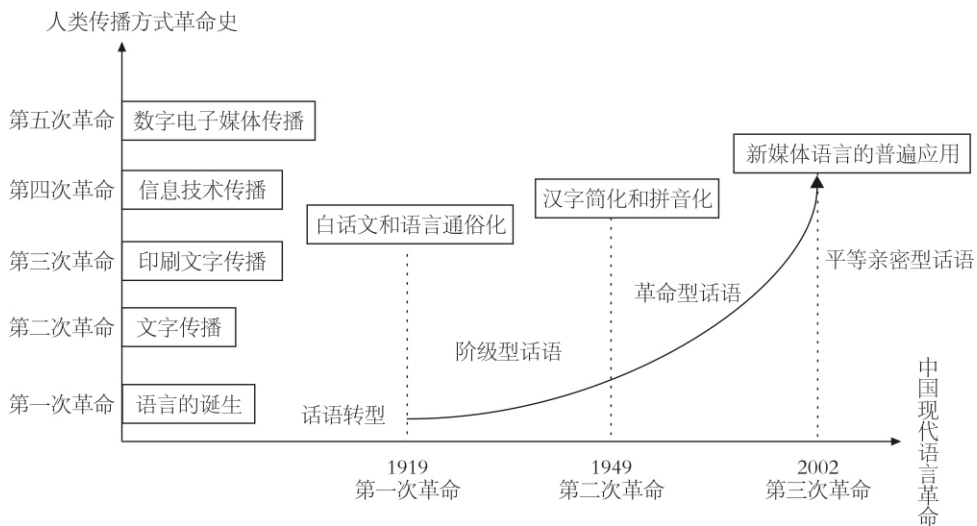


图1 人类传播模式与中国现代语言革命中的话语转型¹

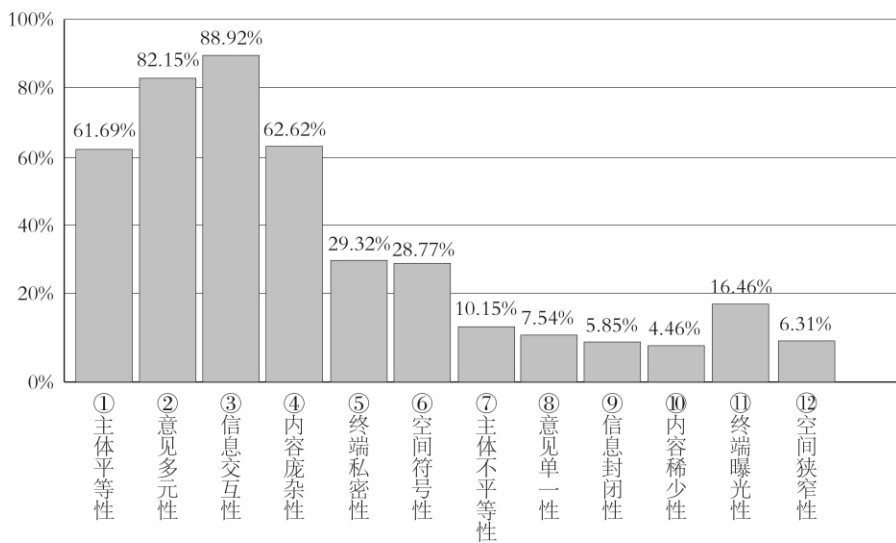


图2 新媒体的特点

播速度很快,传播面很广,发布者和受众可以互动,一些语言形式刚出现的时候大家觉得新奇,于是快速传播,但很快就会变成陈词滥调,语义磨损速度很快,于是就会发生变异,或者创造一些新的语言形式。这些语言形式包括新词、新的用法,或者旧词被赋予新的意义,还有不合常规的缩略语、短时流行的话语形式、字母词和数字词、一些表情态的语气词等等。这些新媒体语言现象在传统平面媒体中很少见到,有一些只在某些圈子里使用,圈外的人不甚了了,例如“白莲花”“圣母白莲花”“玛丽苏”“杰克苏”“非酋”“欧皇”。有些已经传播开来,大家基本上都知道是什么意思,如“直男”“绿茶婊”“白富美”“高富帅”等,尽管这些词语不是人人都会用。我们考察了文学语料、报刊语料、微博语料、科技语料和古汉语语料,可以看到这些词语主要出现在微博语料中。有一些也会出现在报刊语料中,但是数量远远少于微博语料,科技语料

中偶尔会有少量分布,大多是研究性文章所举的例子。(见表1(a)、表1(b)、表2)

	直男	绿茶婊	逼格	屌丝	白富美	高富帅
文学	0	0	0	0	0	0
报刊	2	0	5	58	29	58
微博	1 241	1 915	529	41 282	7 851	9 742
科技	1	0	0	0	0	0
古汉语	0	0	1	0	0	0

表1(a) 新媒体常用词语

一些词素还具有能产性,与其他词素结合构成新词,如“心机婊”“脑残婊”“纯情婊”“红茶婊”“虎牙婊”“布丁婊”“炫富婊”“工作婊”“蜜糖婊”等:

	心机婊	脑残婊	咖啡婊	红茶婊	奶茶婊	工作婊
文学	0	0	0	0	0	0
报刊	0	0	0	0	0	0
微博	3	1	4	11	4	30
科技	0	0	0	0	0	0
古汉语	0	0	0	0	0	0

表1(b) 新媒体常用词语

表1所列举的都是常用名词,这些名词都带有褒贬意义,有些词很不雅,但是使用频率高了之后它本来的意义就淡化了。

	你造吗	有木有	打酱油 ^②	脑子瓦特了	么么哒	吐槽	坑爹
文学	0	0	5	0	0	0	1
报刊	0	5	112	0	2	720	16
微博	540	69 127	4 962	78	20 194	42 722	29 034
科技	0	0	17	0	0	1	0
古汉语	0	0	0	0	0	0	0

表2 新媒体常用词语

表2所列举的都是一些短语,“你造吗”是“你知道吗”的连读音变,“酱紫”是“这样子”的连读音;“表”是“不要”的合音;“有木有”是“有没有”的方言音变;“打酱油”被赋予了新的意义;“脑子瓦特了”是上海话,意为“脑子进水了”,瓦特是英语 water 的音译,“么么”是亲吻动作;“哒”大概跟“好哒”中的“哒”相仿,“吐槽”“坑爹”不须解释了。

	喜大普奔	人艰不拆	十动然拒	不明觉厉	细思极恐	累觉不爱	男默女泪
文学	0	0	0	0	0	0	0
报刊	17	12	7	13	2	2	2
微博	1 413	3 563	421	1 420	337	337	127
科技	0	0	0	0	0	0	0
古汉语	0	0	0	0	0	0	0

表3 新媒体常用缩略格式

表3所列的都是一些新的缩略语。如“喜大普奔”是“喜闻乐见、大快人心、普天同庆、奔走相告”的缩略形式；“人艰不拆”是“人生已经如此的艰难，有些事情就不要拆穿”的缩略形式；“十动然拒”是“十分感动，然而拒绝”的缩略形式；“不明觉厉”是“虽然不明白什么意思，但是觉得很厉害”的缩略形式等等，不一一详解。

	神马都是浮云	你妈叫你回家吃饭	我爸是李刚	哥只是个传说	要么忍，要么残忍
文学	0	0	0	0	0
报刊	2	0	1	0	0
微博	813	30	175	290	85
科技	0	0	0	0	0
古汉语	0	0	0	0	0

表4 新媒体常用句式

还有一些高频的句子也是新媒体常用的，如表4所列。此外还有很多类似的句式，有的流行一阵子就烟消云散了，有的过一段时间又被激活。比如：

- (1) 我从不以强凌弱，我欺负他之前真不知道他比我弱……
- (2) 都想抓住青春的尾巴，可惜青春是只壁虎。
- (3) 没人牵手，我就揣兜。
- (4) 你走你的过街天桥，我过我的地下通道。
- (5) 我的兴趣爱好可分为静态和动态两种，静态就是睡觉，动态就是翻身……
- (6) 一分钱一分货，稀饭吃了不经饿。
- (7) 不能因为咱俩有过节，你就把我当节过。
- (8) 语不惊人死不休。
- (9) 哪里跌倒，哪里爬起……老是在那里跌倒，我怀疑那里有个坑！
- (10) 在海边不要讲笑话，会引起“海笑”的。
- (11) 本是打算搜狗，结果看到猫扑。
- (12) 唉~这人要一没正形，连头痛都是偏的。

此外还有一些字母缩写和数字谐音的语言形式，这些语言形式在微博语料中出现的频率很高。主要原因是用户打字方便，另外还有隐蔽性，有点像“黑话”，圈外的人不懂。当然这些字母词在其他语料中也有，但是意义不同。字母缩略词语在科技文献中很常见，它们都是约定

俗成的,行业性特点很明显。例如:

(13) 九口修返厂研制生产的 ZD-SB 型地膜抢种机,今年全部记设了阻膜轮装五,受到了用户欢迎。(科技语料)

(14) BT 杀虫剂防治棉铃虫应用技术要点吴瑾阿克苏地区农业科技开发中心 EB-82 灭蚜菌剂是中国农科院生物防治研究所研制的新生物制剂,对蚜虫有很强的杀灭作用。(科技语料)

缩略语 SB 和 BT 与新媒体语言中“傻逼”和“变态”的字母缩略词同形。参看表 5:

	GG	SB ^③	BT ^④	Kick	GF ^⑤	BF ^⑥	PMP ^⑦
文学	0	0	0	0	0	0	0
报刊	43	220	331	3	65	31	26
微博	4 853	11 334	31 642	252	2 830	16 378	269
科技	0	76	20	0	12	14	4
古汉语	0	0	0	0	0	0	0

表 5 新媒体常用字母词

	1414 意思意思	9494 就是就是	1314 一生一世	520 我爱你	886 拜拜咯	7456 气死我了	8147 不要生气
文学	0	0	0	0	1	0	0
报刊	0	1	2	?	1	2	0
微博	234	34	2 887	9 946	2 352	25	142
科技	1	0	0	0	1	0	0
古汉语	0	0	0	0	0	0	0

表 6 新媒体常用数字词

表 6 所列是一些数字组合,它们以谐音的方式与一些通用词汇关联。

新媒体语言中有一些语气词使用频率也很高,我们考察了“~ 哒”,发现在文学语料中它几乎都是拟声词,不是语气词。在报刊语料中有拟声词,有方言对音,也有外来语记音,少部分是语气词。在科技语料中“咯咯哒”出现频率很高,这是产品名称,此外还有“咯哒”“叭哒”等拟声词和少量语气词。

	~ 哒
文学	343 ^⑧
报刊	828 ^⑨
微博	46 490 ^⑩
科技	911 ^⑪
古汉语	1 161 ^⑫

表 7(a) 不同语料“哒”的使用频率

为了排除其他用法,我们用“a 哒”“a 滴”“a 咯”和“v 哩”“v 咧”“v 呗”等形式来进行检索,结果如下(这里所列举的“哩”“咧”“呗”都是固有的语气词,有些方言色彩比较明显,但仍然可以观察到在微博语料中它们出现的频率远远高于其他语料。这也许跟语体有关,我们会另文探讨):

	~ 哒	~ 滴(滴)	~ 哈	~ 咯(lo)	~ 哩	~ 咧	~ 呗
文学	0	0	0	2	2 565	476	270
报刊	4	0	0	2	3 014	232	477
微博	5 581	18 289	400	11 436	2 455	3 187	8 283
科技	0	0	0	0	998	287	241
古汉语	0	0	0	0	0	0	0

表 7(b) 新媒体常用语气词

汉语第一人称有“我”“俺”“吾”“余”“某”等形式,在新媒体中,也有用“偶”来代替“我”。

	我	俺	偶 + v
文学	1 991 978	6 324	1 ^⑬
报刊	4 994 543	20 525	0 ^⑭
微博	17 790 374	48 382	13 919 ^⑮
科技	3 932 427	2 783	0 ^⑯
古汉语	2 732 283	50 414	0

表 8 不同语料常用第一人称代词

新媒体语言还有一些语言形式,在其他媒体语言中比较少见,但是在微博语料中大量使用,例如:

	A 歪歪	伤不起	姑凉(姑娘)	我勒个去
文学	1	0	0	0
报刊	0	38	0	0
微博	3 583	31 964	2 876	4 485
科技	3	0	0	0
古汉语	0	0	0	0

表 9 新媒体语言高频语言形式

	有一种 n 叫 n	a 不到哪里去	还 v 呢	爱 v 不 v
文学	3	18	454	38
报刊	31	18	305	72
微博	755	350	3 032	4 019
科技	22	48	132	62
古汉语	0	0	0	0

表 10 新媒体多用构式

新媒体语言还有其他一些独特的语言形式,限于篇幅,这里就不展开讨论了。

二、新媒体语言的修辞学价值

新媒体语言在语言形式和表达方式上有很多创造性的用法,这些用法在修辞方面可以说是很成功的,它增强了语言的表现力。例如:

(15) 世界上最遥远的距离,不是我站在你面前,你却不知道我爱你,而是我们俩一起出门,你去买苹果四代,我去买四袋苹果。(对比/仿拟)

这个句子之所以能够流行,是因为它脱胎于泰戈尔的诗句,这个表达方式是对泰戈尔《飞鸟与鱼》诗句的模仿:

(16) 世界上最遥远的距离,不是生与死,
而是我就站在你面前,你却不知道我爱你。
世界上最遥远的距离,不是我就站在你面前,
你却不知道我爱你,而是爱到痴迷却不能说我爱你,
世界上最遥远的距离,不是我不能说我爱你,而是想你痛彻心脾。
.....

还有一些句子来源于某句歌词,如下面的句子来源于刘若英的《光》,这里的“一下子”和“一辈子”构成对比关系,让人印象深刻:

(17) 你来过一下子,我想念一辈子。(对比)

刘若英在年轻人群体里有很多粉丝,这句话因为歌而流行。

(18) 想问你是不是还记得我名字,
当人海涨潮又退潮几次,
那些年那些事,那一段疯狂热烈浪漫日子,
啊! 恍如隔世——
你来过一下子,我想念一辈子!
.....

也有一些表达是利用双关的修辞手段。例如下面的句子,“放电”是“对自己喜欢的人发出求爱信号”,“来电”指的是“接受求爱信号”和“来电话”,这两个“电”同形不同义:

(19) 不要对哥放电,嫂嫂有来电显示。(双关)

还有一些例子利用跳脱的修辞手段,出人意料。例如:

(20) 姐从来不说人话,姐一直说的是神话。(跳脱)

这是一种修辞手段,意在言外,说话人中间突然转向,改变说话的路径,让听者的思路也跟着跳跃。又如“人生自古谁无死,为了经验怕锤子”“清明时节雨纷纷,多给前任上上坟”,都是突然转变话锋。在修辞手段中,隐喻是无所不在的修辞手法。例如:

(21) 所谓门槛,过去了便是门,过不去就成了槛。(隐喻)

这是一句流行很广的话,有点像“鸡汤”,但是非常有哲理。“门槛”有具体物象和抽象的物象两种语义,是一种双关语义表达。下面的例子也是这样,“两肋插刀”是虚言,“一把刀”是实言。前者是比喻,后者不是。

(22) 我不是不想为你两肋插刀,是因为我只有一把刀。(虚实)

俗话说“姜还是老的辣”,于是有人说出这样的话:

(23) 如果全是老姜,将是一个何等辛辣的社会。(隐喻)

这是一种隐喻的修辞手法,说话人利用既有的语言表达形式中的某个意义范畴,通过隐喻移花接木到另一个意义范畴上。有的时候为了取得幽默效果,利用婉曲和双关手段创造新的说法。例如:

(24) 左脑全是水,右脑全是面粉,不动便罢了,一动全是糨糊。(婉曲/双关)

或者利用语用学的蕴含意义表达某种意义,这种意义要靠听话人自己去联想,去发现这句话的蕴含意义。例如:

(26) 你的丑和你的脸没有关系。(蕴含)

“你的丑和你的脸没有关系”意味着你的丑和你的其他方面有关系,如“你的观念”“你的行为”“你的灵魂”等等。这是一个开放的蕴含关系。有的时候干脆很简单地利用双关手段表达说话人的态度和调侃。例如:

(27) 我这人不太懂音乐,所以时而不靠谱,时而不着调。(双关)

“不靠谱”“不着调”本来的意思是说在音乐造诣上水平低下,但是在现代汉语中只用它们的隐喻意义,“做人做事不可靠,不守规矩”。这里表面上是说不懂音乐,实际上是语带双关,因为“不靠谱”“不着调”本来早就脱离了本义。这个句子因为有上文“我这人不太懂音乐”作背景,把听话人的预期设置到一个语义场中,看似运用了“不靠谱”“不着调”的本来意义,实际上说话人的用意在另外一个语义场,用的还是它们的隐喻意义。

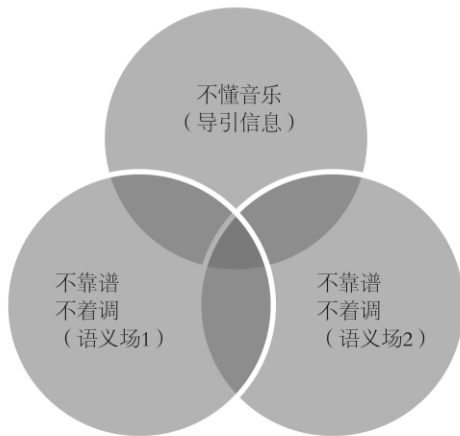


图3 双关涉及不同语义场

三、新媒体语言的社会学价值

通过对新媒体语言的透视,我们可以在社会学意义上了解我们这个社会,了解新媒体的语言使用,了解新媒体的人群。虚拟世界存在着诸多的“圈子”,新媒体语言在不同的圈子里有不同的潮流,这些潮流有的时候可以反映出一个社会群体的集体焦虑、集体无意识、集体愤怒,也可以反映出不同社会群体的价值取向、社会心理。每个圈子都有自己的标签和自己的圈子文化,圈内人在这里可以找到知己,获得认同感。这些圈子大小有很多,粗略地看有一些(有些圈子里还可以细分成不同的更小的圈子):

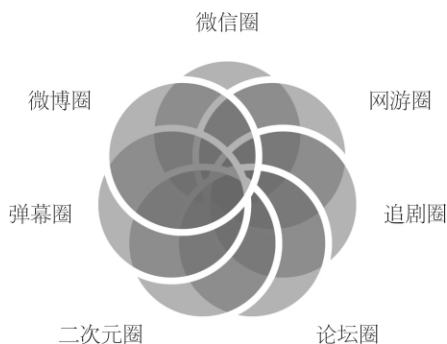


图4 新媒体社交圈示意图

各个圈子彼此之间也有交叉,并不是泾渭分明的。有交集的地方所创造出来的语言形式往往可以成为社会通用语的一部分。下面我们分几个主题来讨论新媒体语言的社会学价值。

1. “标题党”与媒体文化

我们看网络新闻经常被“标题党”忽悠,点开之后才发现内容与标题风马牛不相及。这在新媒体圈子里几乎成了当下约定俗成的现象,大家对此颇有异议,有违语言传播表义准确明晰的基本规律,但是却无法根除。根据中国互联网络信息中心(CNNIC)2018年1月31日发布的第41次《中国互联网络发展状况统计报告》,中国网民规模已经达到7.72亿,手机网民占比达到97.5%,人数达到7.53亿。这意味着新媒体已经成为人们获取信息的主要渠道。在海量信息中,如何才能抓住读者的眼球就成了很多信息发布者的头等大事。标题显然是非常重要的,他们在标题上做文章,主要手段有夸大其词、制造悬念、刺激暗示、激将法等等。例如:“史上最牛……”“全中国都在看……”“大家注意啦,不看要吃亏啦”“中国股市要有重大变化”“刚刚传来的消息:特朗普即将遭到弹劾!”等等,标题都有些危言耸听,看内容则严重失实。又比如“震惊!管理层出游竟然包下一辆火车!上市公司员工爆出惊天内幕”“最新重磅:崔永元冒死实名举报的竟然是他”“他们找到了治疗癌症的新方法”“太可怕了!女子睡前做了这件事,竟然导致不孕不育”“为什么说住酒店最好不要关厕所灯?酒店服务员偷偷告诉你原因”“欧洲美女为什么不愿意嫁中国男人?”这里边既有故意卖关子、制造悬念,也有一些暗示。有的时候还会用到激将法,例如“是中国人的转发”“是男人一定要知道”。关于标题党现象,很多人已经提出批评,例如贾晓红(2018),王诗瑶、石晋阳(2018)等等,也有人对新

闻标题进行比较研究,发现了新媒体与传统媒体的差异(颜晓尹 2018)。吴振东、王艳称这种现象是“标题党思维”,这是很有道理的。这种标题党思维已经渗透到新媒体语言的骨髓中。我们从微博和报刊中可以看到标题党思维的确已蔚然成风。下面是 BCC 语料库中的表达反映出来的标题党思维:

	史上最	吓死	大尺度	惊呆了	曝光	狂 V
微博	17 430	15 265	1 701	25 172	22 654	15 465
报刊	1 362	158	131	818	6 336	42 945

表 11 新媒体表达反映出来的标题党思维

2. 网络语言暴力与社会戾气

新媒体时代也是自媒体时代,在虚拟社会里,语言暴力现象比较突出,这可以从一个侧面看出日益严重的社会戾气。新媒体的一个特点就是某些极端的表述很容易形成某种趋势,跟风的人越来越多,公正的、理性的言论往往被语言暴力淹没。有人注意到了自媒体语言暴力的治理问题,指出很多社交平台设置了暴力语言过滤软件,如“垃圾”“去死吧”“蠢猪”“去你妈”“智障”“人渣”“败类”“婊子”等暴力词汇通过技术手段可以屏蔽,但是有很多语言暴力现象是以隐蔽的方式存在的,例如“卧槽”“我擦”“我去”“买了表”“二百九(二百五+三八+二)”等等。这些伪装起来的语言形式,大部分是詈骂语或不雅的话,也有一些是为了避讳或应对审核(如“毛衣战”是“贸易战”的伪装形式)。我们从语料库中也可以看到,其他语料出现这些语言暴力表达的频率几乎为零。例如:

	猫了个咪	买了个表	我去年买了个表	草泥马	牛叉	特么
文学	0	0	0	0	0	0
报刊	0	0	0	1	0	0
微博	128	3 373	2 160	15 897	1 342	14 606
科技	0	0	0	0	0	0
古汉语	0	0	0	0	0	0

表 12 新媒体经过伪装的语言暴力现象

“垃圾”“去死”“蠢猪”“去你妈”“智障”“人渣”“尼玛”“婊子”这一类词汇并非是新媒体特有的,然而我们可以看到,它们在新媒体中出现的频率远远高于其他媒体。

	人渣	婊子	傻逼	尼玛	去死吧	智障
文学	36	1 055	4	51	81	17
报刊	12	51	0	1127	7	268
微博	4 270	5 139	58 585	94 566	4 582	1 965
科技	30	97	4	242	25	219
古汉语	0	507	0	2 191	4	1 283

表 13 不同语料中语言暴力现象统计比较

以“尼玛”为例,在微博语料中几乎都是詈骂语(“你妈”的变身),且出现频率极高,而在其他语料中基本上是人名,而“尼玛”在藏语中是“太阳”的意思。

3. 新媒体语言所反映的社会价值取向

新媒体语言出现的这些现象其实都是社会现象在虚拟空间的真实反映,它反映了某些社会群体的社会价值取向。例如“佛系”一词的流行,反映的是某些人的生活态度。请看下面一段文字:

(28) 这么高的人气,或许是因为现代社会压力太大,许多民众其实都想放下脚步,放空身心,享受片刻无欲无求的宁静感。这与本周大陆突然火爆的“佛系 90 后”概念其实有着类似的内涵。“佛系员工”(指交代的任务从不推辞,但也绝不会给自己揽工作,工作质量永远保证合格,但从不会有惊喜)。“佛系恋爱”“佛系点赞”……跟台湾一样,大陆一线城市年轻人工作、学习、生活压力很大,已经觉得很辛苦,所以推崇“一切随缘”。一方面,这是好事,不争不抢,随性豁达。另一方面,有时也有面对竞争和压力,消极怠惰之嫌。好与不好,存乎一心。(人民日报海外版 2017. 12. 23)

这段文字是对“佛系”一词最好的注释。还有一些词汇带有歧视性和侮辱性,如“打工仔”“死基佬”“变态佬”“矮穷挫”“死八婆”“剩女”“玻璃”等,这都反映出一些不健康的社会价值观念。

4. 新媒体语言的嬗变与大众心理

新媒体的语言也是在不断变化的。这种嬗变也从一个侧面反映出社会面貌的变化和世态人心的变化。例如 BCC 语料库中关于“走后门”一词的历史检索可以看出这个词的热度是在不断变化的^⑩:

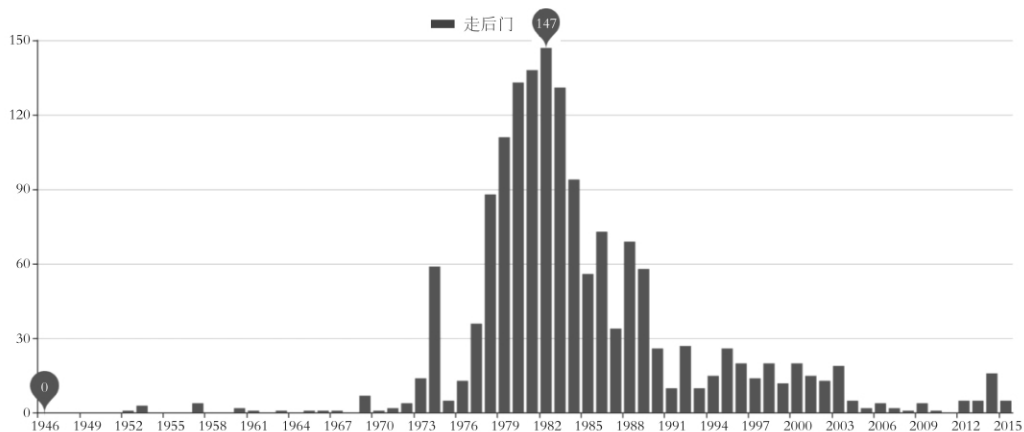


图5 “走后门”一词的热度变迁

我们也可以看到,“吃软饭”一词主要在微博语料中出现,而“吃闲饭”则各个语料都有分布:

	吃软饭	吃闲饭
文学	7	31
报刊	4	321
微博	474	113
科技	7	76
古汉语	7	24

表 14 “吃软饭”与“吃闲饭”比较

在新媒体语言中,有些语言形式出现频率很高,如“帅哥”“美女”,但是我们也看到它们出现的时间段集中在新媒体出现之后。

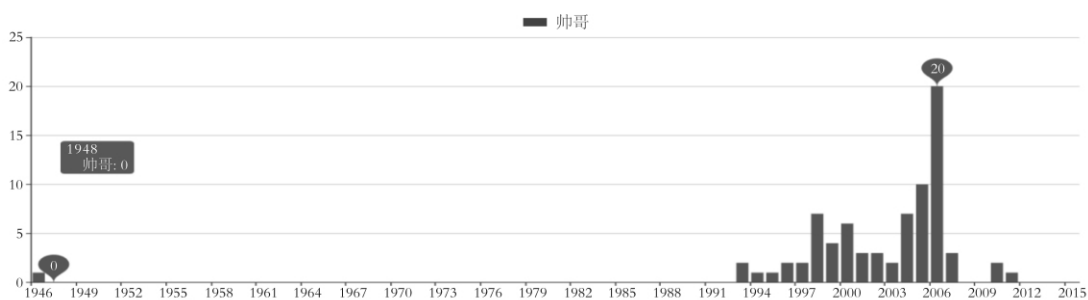


图 6 “帅哥”一词历时分布

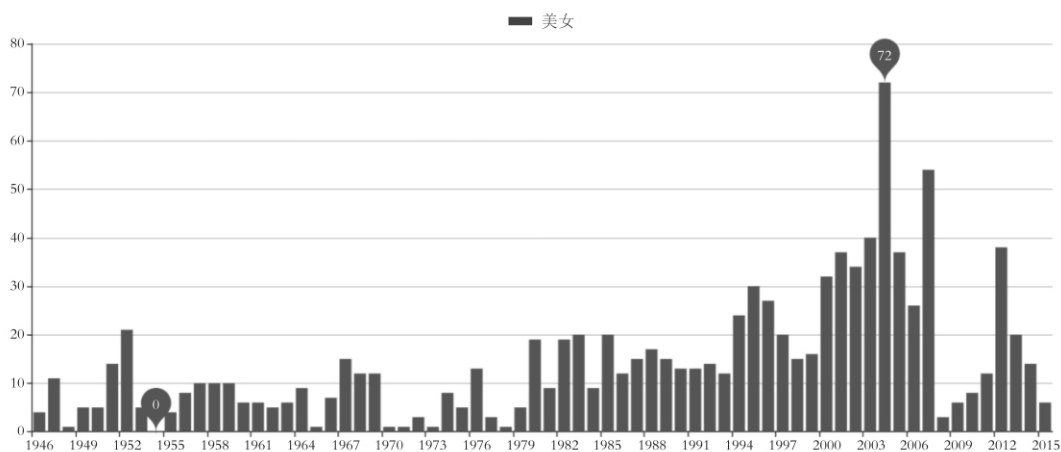


图 7 “美女”一词历时分布

由于 BCC 语料库的历时检索主要局限在报刊语料,所以这个图表不能反映“帅哥”“美女”历时分布情况的完整面貌。称谓问题在中国社会是一个敏感问题,称呼错了会给自己带来很多麻烦。有一段时间社会上凡是男人都称“老板”,凡是女人都称“小姐”。现在这两个称谓形式已经式微:

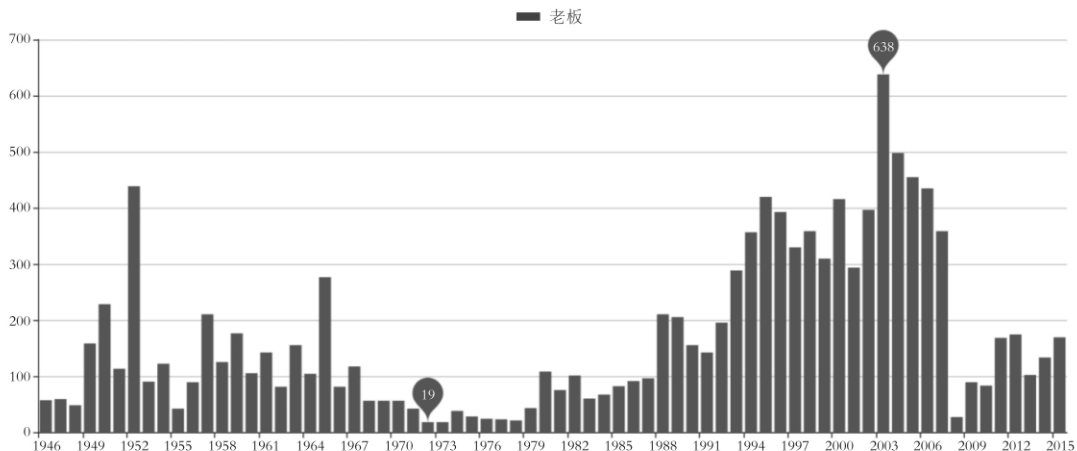


图8 “老板”一词的历时变迁

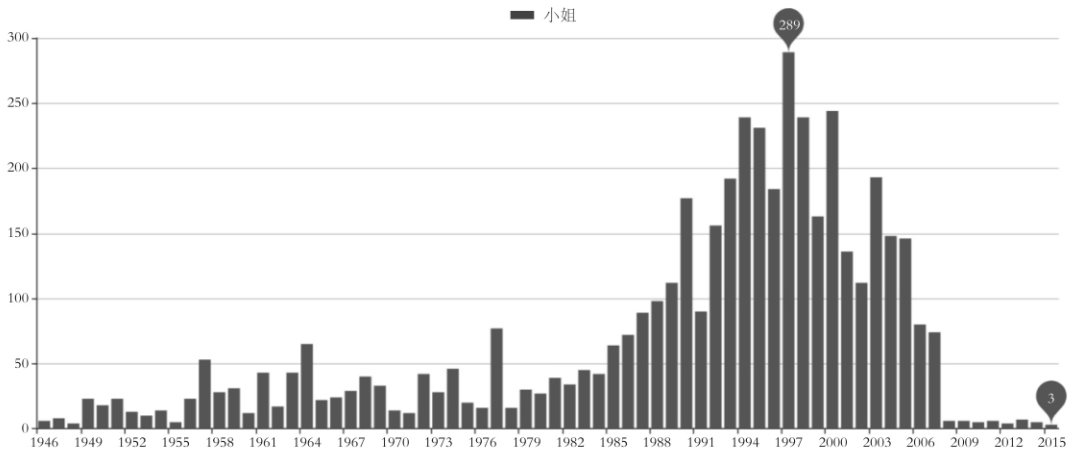


图9 “小姐”一词的历时变迁

在上个世纪末“十亿人民九亿商,还有一亿待开张”的年代,“老板”是个热度很高的称谓,“小姐”也是个热度很高的称谓。现在这两个词开始变冷。从这种变化中我们可以窥见世态人心的变化。

注 释

- ① BCC 汉语语料库,总字数约 150 亿字,包括:报刊(20 亿)、文学(30 亿)、微博(30 亿)、科技(30 亿)、综合(10 亿)和古汉语(20 亿)等多领域语料,是可以全面反映当今社会语言生活的大规模语料库。
- ② 微博语料和报刊语料的“打酱油”是“路人”“不相干的人”“凑热闹的人”,文学语料和科技语料的“打酱油”是“用瓶子买酱油”的意思,它们是“打酱油”的原意。
- ③ 报刊语料和微博语料 SB 是“傻逼”的简拼,科技语料不是。
- ④ 微博语料的 BT 是“变态”的简拼,报刊语料有一部分是“变态”的简拼,科技语料没有“变态”的用法。
- ⑤ 微博语料 GF 是英文 girlfriend 的简写,报刊语料也有这种用法,科技语料的 GF 没有这种意思。

- ⑥ 微博语料 BF 是英文 boyfriend 的简写,报刊语料也有这种用法,但很少,科技语料的 GF 没有这种意思。
- ⑦ 微博语料 PMP 是“拍马屁”的简拼,其他语料没有这种用法。
- ⑧ 文学语料中的“哒”大多是拟声词,不是语气词。
- ⑨ 这里的“哒”有一部分是拟声词,一部分是语气词。
- ⑩ 只有少量的拟声词,大部分是语气词。
- ⑪ 这里的“哒”有拟声词,有“哪哒儿”中的“哒”,还有一些是外来语记音,少部分是语气词。
- ⑫ 几乎都是拟声词。
- ⑬ 检索文学语料共得到 388 个“偶 v”的例子,只有 1 例是人称代词的用法。
- ⑭ 检索报刊语料共得到 2 284 个“偶 v”的例子,我们考察了 1-1 000 例,没有一个是人称代词的用法。
- ⑮ 这里的例子绝大部分是第一人称用法。
- ⑯ 检索报刊语料共得到 11 464 个“偶 v”的例子,我们考察了 1-1 000 例,没有一个是人称代词的用法。
- ⑰ BCC 语料库的历史检索请参看荀恩东、饶高琦、谢佳莉、黄志斌 2015。

参考文献

- 谷云峰 2018 网络流行语言看青少年的价值观特点,《教育现代化》第 26 期。
- 霍慧娜 2017 媒体融合背景下新媒体语言特点研究,《新媒体研究》第 20 期。
- 贾晓虹 2018 新媒体“标题党”乱象及治理,《新媒体研究》第 7 期。
- 罗 譞 2018 网络语言的价值观引导与治理,《媒介经营与管理》第 17 期。
- 李 琼 2016 新媒体时代网络新闻报道的语言特点,《今媒体》第 11 期。
- 李 雪 2016 新媒体对大学生价值观的影响及其对策研究,广西师范学院硕士学位论文。
- 尚 伟 2017 新媒体语言的发展及其规范,《北华大学学报》(社会科学版)第 4 期。
- 施麟麒 2018 政务新媒体的社会语境与语言选择,《语言文字应用》第 1 期。
- 王建华 2017 《政务新媒体话语应用与传播研究》,上海交通大学出版社。
- 王建华 2018 政务新媒体语言应用,《语言文字应用》第 1 期。
- 王诗瑶、石晋阳 2018 纸媒公众号“标题党”现象与受众认同分析,《青年记者》7 月下。
- 荀恩东、饶高琦、谢佳莉、黄志斌 2015 现代汉语词汇历时检索系统的建设与应用,《中文信息学报》第 29 卷第 3 期。
- 吴振东、王 艳 2018 噱头的诱惑与媒介素养——新媒体时代语境下的“标题党”思维分析,《新媒体研究》第 9 期。
- 颜晓尹 2018 新媒体和传统纸媒新闻标题语言对比——以《人民日报》为例,上海外国语大学硕士学位论文。
- 赵 莉 2014 新媒体科学传播亲和力的话语建构研究,中国科学技术大学博士学位论文。
- 赵 雪、曹彦男 2012 新媒体语言研究十年,《理论与现代化》第 4 期。

A Corpus-based Study of New Media Language

Cui Xiliang

Abstract: The new media language, with its own characteristics, have been studied frequently. However, the corpus-based researches are still rare. Therefore, this paper, based on the BCC corpus, examines the various linguistic phenomena of new media and analyzes its linguistic, rhetorical and sociological values.

Keywords: new media language, corpus-based, linguistics, rhetoric, sociology