

《现代汉语词典》词类标注变化的统计与评析

尹洪波

【提要】《现代汉语词典₆》在《现代汉语词典₅》的基础上,对词类标注做了一定的修订。经全面统计发现,这些修订包括调、改、增、删四个方面,涉及9种词类,但名、动、形的变化频次,尤其是名和动的变化频次,占绝对优势。这有力地印证了沈家煊先生近年来从不同角度所论证的观点:汉语里名和动之间是“名动包含”模式,即汉语的动词是名词所包容的一个次类。《现代汉语词典₆》词类标注的变化,使我们更为真切地感受到汉语词类标注之难,主要在于两条界线:一是名、动、形三大实词之间的界线,二是词与非词(语素和短语)之间的界线。经本次修订,《现代汉语词典》词类标注的准确性得到了较大提升,当然也尚存进一步提高的空间。

【关键词】《现代汉语词典》 词类标注 统计评析

【中图分类号】H164 **【文献标识码】**A **【文章编号】**1000-2952(2016)03-0108-07

从1960年开始发行试印本以来,《现代汉语词典》(以下简称《现汉》)已经跨越两个世纪,历时五十多个春秋。2005年《现汉》(第5版)(以下简称《现汉₅》),首次做了全面的词类标注。七年后,《现汉》(第6版)(以下简称《现汉₆》)问世。在《现汉₅》的基础上,《现汉₆》主要做了9个方面的修订,其中第六个方面是:“复查了词类标注,在保持原有词类标注体系的基础上,对少数词的词类标注做了修订。”(见《现汉₆》“第6版说明”)本文首先全面统计《现汉₆》词类标注的变化,然后加以评析。

为节省篇幅,用“名、动、形、副、介、连、数、量、拟”等分别表示“名词、动词、形容词、副词、介词、连词、数词、量词、拟

声词”,“未”字表示未标注词类。

一、词类标注变化统计

经逐一核对统计,与《现汉₅》相比,《现汉₆》词类标注的变化,可概括为调、改、增、删四个方面,详情如下:

(一) 调

《现汉₆》调整了一些多义词的义项排序,这些义项词性可能相同,也可能不同。如果不同,义项排列次序的调整,就同时也带来义项词性排序的变化。因此,这里的“调”,是指兼类词不同词性的义项排列次序的调整(有的义项调整不涉及所标注词类的排序的变化,不作统计)。例如:安慰,《现汉₅》标注为{1形2

动),《现汉》把次序调整为{2动1形}(数字1、2表示义项的序号,以下准此)。经统计,兼类词不同词性排序的调整,共有以下几类:

第一,“名—动”类:(1)“名动→动名”的,7个:摆¹、比喻、观瞻、擒拿、通关、赢利、重奖;(2)“动名→名动”的,2个:出品、秃头。第二,“形名→名形”类,5个:二把刀、二五眼、傻帽儿、土产、自由。第三,“副名→名副”类,6个:梯次、梯度、万难、沿路、沿途、一头。第四,“形动→动形”类,6个:安慰、绕远儿、润、透、长(zhǎng)、调和。第五,“副动→动副”类,9个:保不住、怪不得、独身、难怪、现、一气、怨不得、在先、着意。第六,“副形→形副”类,6个:大不了、交关、累累、任情、微微、早。第七,“介动→动介”类,4个:根据、就、通过、依据。第八,其他,5个:拐¹、样、第³、稀、传。

(二) 改

指词类标注的改变,原来标为甲类词,现在标为乙类词。下面统计列举时,对单义词,直接列出该词。对多义词,则在该词后列出词类标注发生变化的那个义项的序号。词类改标的情况共有九类:

第一,“名—动”类:(1)“名→动”的,26个:常行军、传达2、对话1、傅科摆、歌舞、归纳2、滑冰1、滑雪2、科普、跨栏、类比1、论证1、判断1、热传导、热辐射、热膨胀、赛艇1、商检、射击2、射箭2、实践2、挺举、网恋、想象1、演绎1、怨怼。(2)“动→名”的,4个:干打垒1、肝硬变、肝硬化、空战。第二,“形—动”类:(1)“动→形”的,28个:懊恼、憋气、不才1、不佞1、得人儿、动摇1、独立3、寡合、过心2、合不来、合得来、横3、狐媚、坏3、口臭1、赖皮1(原2动)、雷同、闹心2、逆耳、怕事、怕羞、清¹7、贪杯、贪嘴、无情2、心爱、淫乱1、应试2。(2)“形→动”的,9个:重沓、敌对、费事、固定1、恼怒1、疑似、御用1、主打、主导1。第三,“名—形”类:(1)“形→名”的,6个:动态3、红热、上上、实木、所属2、下下1。

(2)“名→形”的,9个:独门2、孤身、海量3、紧身、垃圾2、廉价、特务(tèwù)、微量、一站式。第四,“形—副”类的:(1)“副→形”的,2个:大约1、亲身。(2)“形→副”的,3个:个别1、悍然、浑然2。第五,“动—副”类的:(1)“动→副”的,2个:不止2、协力。(2)“副→动”的,2个:可能3、席地。第六,“名—副”类的:(1)“名→副”的,1个:光景4。(2)“副→名”的,3个:实地1、私下2、独身。第七,“名→量”类,2个:更2、歇4。第八,原为其他词类,改为“代”的:(1)“名→代”的,1个:奴家。(2)“副→代”的,3个:多6(原多²之1、3)、多么1、好15。(3)“连→代”的,1个:首先2。(4)“数→代”的,1个:几1。第九,原为其他词类,改为“介”的:(1)“连→介”的,3个:比及、等²2、等到。(2)“动→介”的,2个:临到2、有关2。

(三) 增

有两类情况:一是单纯增标词类,即某个义项原来未标词类,现在增加标注;一是增加义项,同时标注词类。为节省篇幅,下文列举时,对只有一个义项的词,直接列出该词;对于多于一个义项的,则在该词后面写出增标词类的义项的序号。增标两类的只有4个:单身1、3,驾5、6,纠结2、3,绅士2、3。绝大多数是增标一种词类的,下面予以重点讨论。

1. 单纯增标词类的,总共68个,又分为“+名”“+动”“+形”“+其他”四小类(+表示增加):(1)“+名”的31个:叉(chā),钗,筹1,锤1,大军2,第一把手,杠1、2、4,胳膊腕子,胳膊肘子,勾²,狗吃屎,怪5,满堂彩,杪1,妮,皮5,炔,纱2,痧,上来2,上下2,首位,榻,筒1,突5,铤,未²,新纪元,洋2,糟1,蝎。(2)“+动”的17个:按²2、不敢当、不过意、缚1、哈²、轮渡1、满负荷、撮¹2、酬¹2、双肩挑、綯2、唾2、养5、一锅端1、2,一锅煮,一勺烩,拐2。(3)“+形”的9个:不得了1、2、3,不是味儿1、2、3,不像话1、2,不要紧、诚1、痴4、悄没声

儿、揪、死乞白赖。(4)“+其他”的6个：“+叹”的1个：阿门。“+副”的5个：会²6，窃4，素6，徒8、9，委⁵。

2. 增加义项，同时增标词类的，总共231个：

第一，增加“名”类义项的，共74个，几乎都是兼类词。据兼类情况，可细分为四种：

(1)“名动”兼类的，50个：备料2、布局3、采编2、重叠2、丰乳2、丰胸2、红脸3、际遇2、检讨2、忌讳3、决口2、开口¹3、开山4、库藏2、烙花2、雷3、裸体2、偶感2、配料2、签名2、签字2、晒图2、失误2、失着1、替补2、挑战4、拖腔2、网4、卧底2、信仰2、遗患2、疑惑2、疑虑2、引港2、援手2、正点2、正误1、主持3、主创2、主刀2、主讲2、主使2、转弯3、撞锁2、转弯3、撞锁2、自称2、自信2；编审3、梦3。这些词，前48个原来都是动词。后2个，原来就是名动兼类词。(2)“名形”兼类的，10个：草灰1、骨感1、虎气1、近情2、铁血1、铁杆2、危险2、血腥1、紫花1；高层4。这10个词，前9个原来都是形容词。最后1个，本来就是名形兼类词。(3)名量兼类，2个：台¹6、摊子2。上述三种都是兼类性增加，另外还有少量非兼类性增加，12个：粗工2、恶人2、蚰2、后¹6、绝招儿1、栏2、梁¹5、署¹2、岫3、满堂红2、镗2、生³3。

第二，增加“动”类义项的，共77个，其中少数为非兼类性增加，绝大多数为兼类性增加。兼类性增加可细分为三类：(1)“名动”兼类的，43个。本来是名词或动词的，35个：包桌1，被6，吃货1、2，垫话1，讹传1，恶作剧1，儿戏1，规律2，积分¹1，寄籍1，孑遗1，劲歌2，劲舞2，具文1，开口¹2，力证2，昵称2，陪客1，枪战1，人治2，失地3，谈吐1，搪便2，挑战3、听闻2，投向1，外感1，万事通1，委曲3，蜗居2，限额1，谐声1，行商2，意图2，铸币1。本来是名动兼类词的，8个：感冒3、计2、加餐2、来电3、拍4、起笔3、认识2、造¹3。(2)“形动”兼类的，18个。本来是形容词或动词的，16个：把牢1、2，黑6、

7，挥斥2、3，均2，啰嗦3，迷糊2，平衡3，强健2，软7，松懈4，调协2，贴边¹1，歪3，哑4，淫乱2，忠实3。本来就是形动兼类的，2个：憋3、废4。(3)“动副”兼类的，2个：接连1、算是1。非兼类性增加的，14个：得9、调¹3、絨2、絨2、捏2、排²2、怂2、弹5、探3、调4、挺4、洗手3、宅2、作7。

第三，增加“形”类义项的，共45个：

(1)“形名”兼类的，29个：壁挂2、菜5、差额2、长期2，潮¹4，粗线条2、3，短期2，二百五2，盖¹7，过6，牛3，权威3，热心肠2，荣誉2，亮7，软3、8，山寨3、4，数码2，数位2，数字4，饕餮4，痛痒1，问题5，下乘2，现代2，直肠子1，自信3，轴5，专业4。(2)“形动”兼类的，10个：没谱儿2、配9、失败3、随便4、贪4、外露2、有谱儿2、有心2、整齐5、执行2。(3)“形副”兼类的，2个：多方1、硬6。前一个本来是副词，后一个本来就是形副兼类词。(4)“形拟”兼类的，1个：铮铮2。以上都是兼类性增加，还有3个非兼类性的：暗4、浅7、早熟2。

第四，增加“副”类义项的，共18个。其中形副、动副兼类的略微多一些，不再细分：保险4，比肩2，纯4，独5，少6，保准3，附带2，过7，接着2，指定2，实3，常2，通常2，强力3，全力2，多一半2，要不3，特¹4、5。

第五，增加“连”类义项的，共5个，不再细分：还是5、和²4、回头5、凭5、叵测3。

第六，增加“介”类义项的，共6个。这些词原来都有动词义项，不再细分：给7、8，叫8，连4，拿10，让9，用6。

第七，增加“量”类义项的，共6个。这些词原来都有名词义项，不再细分：捆子2、链2、畦2、摊4、摊子3、桶2。

(四) 删

有两种情况：

1. 直接删除某义项原来标注的词类的，共38个：(1)“一名”的，21个：广²1、号¹4、黑3、后¹5、环3、涝2、墨¹9、母4、上声、舌3、

兽¹、署¹、垮、胃²、窝⁵、弦⁵、解²、榭、型²、志愿兵役制、訇。(2)“一动”的，15个：不经意，筹²，活受罪，计⁵、6，零七碎八²，乱弹琴，偏利共栖，偏利共生，破天荒，切⁵，捌，畏¹，挟¹，重⁵，作²、3。(3)“一形”的，2个：呱呱叫、乱七八糟、零七碎八¹。

2. 删除某个义项，从而导致该义项的词性标注也随之删除的，共41个：(1)删“名”类义项的，26个，又可分为三小类：其一，名动兼类被删“名”的，14个：表示³、表现²、传承²、代步²、地震²、懂²、公祭²、胡说²、记载²、警告³、救护²、譬²、清唱¹、赛车³；其二，名形兼类被删“名”的，8个：不平²、4，赤忱²，大²，大气(dàqì)¹，黑³，久²，快²，脛²；其三，名量兼类被删“名”的，2个：方里¹、方丈¹。以上都是兼类词被删除“名”的。此外，还有非兼类词被删“名”的，有2个：后¹、署¹。(2)删除“动”类义项的，8个，又分两小类：其一，名动兼类被删“动”的，6个：爱好³、便饭²、导向²、教授(jiàoshòu)²、新交¹、盈利¹；其二，形动兼类、副动兼类被删“动”的各1个：沮丧²、无何²。(3)删除“形”类义项的，3个：光大²、兼任²、皮肤²。(4)删除其他词类义项的，4个：才²、容许²、及至²、场⁸。

以上是《现汉》词性标注修订的所有信息，下面对这些变化做一些简要的评析。

二、词类标注变化评析

我们先按前文“调、改、增、删”的次序，对词类标注变化逐项评析，然后做一个总结。

(一)“调”类变化评析

前面已经说过，《现汉》调整了部分多义项的义项排序，从而导致不同义项词性排序的变化。“调”类变化共涉及50个词，名、动、形、副、介、连、数、量8种词类。这些调整“更加注重词义发展顺序”(张博2013)，^①也符合吕叔湘(1958)^②当初所制定的义项排列原则：“词义分项排列的先后，基本的在前，引申的在

后；一般的在前，特殊的在后；具体的在前，抽象的在后。”比如把“介动”排序调整为“动介”，符合汉语词类发展演变的事实，因为汉语里的介词几乎全都是从动词语法化而来的。又如把“形名”调整为“名形”，把“副名”调整为“名副”，这是很有道理的：根据 Bernd Heine and Tania Kuteva (2007)，^③从名词演变出形容词、副词，在人类语言中具有一定的普遍性。

当然也有可以进一步讨论的地方。比如“长(zhǎng)”原来是形容词义项在前，动词义项在后，现在反之。到底哪一种排序符合汉语的事实，恐怕尚需研究而定。因为，目前我们尚无足够的证据支持从动词到形容词的演变路径，而古汉语中许多形容词可以用作动词倒是不争的事实。

(二)“改”类变化评析

词类改标共涉及108个词，名、动、形、副、代、介、连、量、数9种词类；其中名、动、形三大实词类占大多数：“名—动”(30个)，“形—动”(37个)和“名—形”(15个)三类共计82个，占总数的76%；另外，“名—动”类改标的绝大多数是“名→动”的(26个)，占87%；“形—动”类改标的绝大多数是“动→形”的(28个)，占76%；“名—形”类改标的，“形→名”的(6个)和“名→形”的(9个)相当。

这些修订绝大多数值得肯定。比如“滑冰、滑雪、跨栏、赛艇、射击、射箭”等运动项目类词语及其他一些名动兼类词，《现汉》一律都改标为“动”，不再按名动兼类词处理。这些词既表示某种运动，同时又是某种运动项目的名称，这大概是最初标为名动兼类的理由。这样

① 张博：《〈现代汉语词典〉第6版释义修订的类型及特征》，《辞书研究》2013年第2期。

② 吕叔湘：《〈现代汉语词典〉编写细则》，商务印书馆1958年版。

③ Bernd Heine and Tania Kuteva, *The Genesis of Grammar*. Oxford: Oxford University Press, 2007, pp. 57—120.

标注显然是受了印欧语词类观的影响：印欧语里的动词用于指称时要“名词化”，例如 publish → publishing/publication。但是，汉语没有形态变化，动词用于指称时，不需要任何形态变化。沈家煊（2009：1~12）^①指出，“汉语陈述语用做指称语是‘构成关系’”，“动词用作指称语，是指一个被视为实体的活动，是一种‘本体隐喻’”，“对于中国人来说，一个活动就是一个实体，所以不需要改变词形”。由名动兼类改标为“动”的，除一批运动项目类词语外，还有“归纳、类比、论证、判断、实践、想象、演绎”等词语。如果我们承认汉语里的动词不需要改变词形，可直接用于指称，这些词便可以名正言顺地标为“动”，无需繁琐地标为名动兼类词。这一修订，算得上是词类标注的一次重大改进，不仅体现了修订者对汉语名词、动词认识的深化，也很好地贯彻了词类划分的简约原则。

然而也有值得商榷的地方。比如“肝硬变、肝硬化、空战”原标“动”，现标“名”。本文认为还是标注为“动”更合适些。“傅科摆”（一种仪器）原标“名”，现标“动”，“恼怒”的义项1原为“形”，现标“动”，这显然欠妥。

（三）“增”类变化评析

“增”类变化共涉及名、动、形、副、连、量、介、叹8种词类，303词。其中增加两种词类的极少（4个），绝大多数（299个）是增加一种词类的。在只增加一种词类的299词中：单纯增加词性标注的只有68个，占23%；大多数（231）是因增加义项而导致词类标注的变化，占77%。无论是单纯增标词类，还是增加义项，其中都是名、动、形占大多数：占前者总数（68）的84%，占后者总数（231）的85%。

在“调改增删”四类变化中，“增”类最为复杂，所涉及的词语数量最多。“增”类的修订，绝大多数是值得肯定的。例如“皮”，原来认为义项5（某些薄片状的东西）是非成词语素，未标注词类，现在认为是成词语素，增标了“名”。又如：满堂彩、不得了、不敢当、不

过意、不是味儿、不像话、不要紧、满负荷、双肩挑、悄没声儿、一锅煮、一勺烩、第一把手、胳膊腕子、胳膊肘子、死乞白赖，这些多字条目，过去认为是短语，未标词类，现在承认它们是词，增标词类。又如：采编、决口、库藏、配料等词语，属于典型的名动兼类词，过去只有一个动词义项，现在增补了一个名词义项；而“恶作剧、劲歌、劲舞、枪战”等典型的名动兼类词，过去只有一个名词义项，现在则增补了一个动词义项。再如：迷糊、平衡、强健、忠实，属于典型的形动兼类词，过去只有形容词义项，现在增补了动词义项。再如：菜、权威、荣誉、现代，属于形名兼类词，过去只有名词义项，现在增补了形容词义项。当然也存在可以讨论的地方。例如：新纪元，过去看作短语，现在看作是词。本文认为还是看作短语更好些，理由是：“新纪元”显然是由“新”和“纪元”这两个词组合而成的，二者之间的关系仍然比较松散，可以插入“的”扩展为“新的纪元”。又如：短期、长期，过去都标为名词，增加了形容词义项，看做名形兼类词。笔者认为，这有悖于简约性原则。因为新增加的形容词义项的意义和原来的名词义项的意义没有本质的不同，不宜看作兼类词。

（四）“删”类变化评析

“删”类修订所涉及的词语共计79个，其中只删除原来所标词类的38个，删除某个义项而该义项的词性标注也随之删除的41个。所涉及的词类共有名、动、形、副、连、量六类，其中涉及频次最多是：名（47）、动（23）、形（5），总计74次，占总次数（79）的94%，而名动又占绝对优势，占总数的89%。

“删”类的修订，有些是对过去明显失误的修正。例如：呱呱叫、乱七八糟、零七碎八，前一个是熟语，后两个是成语。根据《现汉》的词类标注体例，都不应标注，过去都做了标注，属于明显的失误。其他绝大多数修订，值得肯定。例如：表示、表现、传承、地震，过

^① 沈家煊：《我看汉语的词类》，《语言科学》2009年第1期。

去都标为名动兼类，现在则一律删除其名词义项，不再看作名动兼类词。这样处理的优越性和合理之处，前面“‘改’类变化评析”部分已有详细论述，此处不赘。

以上分别对“调、改、增、删”四个方面的修订变化分别做了评析，下面做一个总结。

《现汉₆》对词类标注的修订，总共涉及约540个词语和名、动、形、副、代、介、连、数、量9种词类。这些修订，与过去相比有明显的进步和提高，同时，也突显了汉语词类标注的一些老大难问题：首先，是词与非词的分界问题。汉语不像印欧语，能够自然分词，有时词的判定和确认是件十分困难的事。在“语素—词—短语”这三级单位中，词与其下一级单位语素存在划界问题，与其上一级单位短语也存在划界问题。正如吕叔湘（1979：17）^①所言：“词在两头都有划界问题：一头是如何区别单独成词的语素和单独不成词的语素；另一头是如何决定什么样的语素组合只是一个词，什么样的语素组合构成一个短语。”在词典词性标注具体操作中，对于单字条目，需要判断是否成词，对于多字条目需要判断是词还是短语。这都是十分繁难的事情。对于单字条目是否成词，还涉及到古代汉语和现代汉语的问题。古汉语里，大多数词都是单音词。有些单字，从古代汉语的角度看是词，从现代汉语的角度看可能不是词，而是语素。然而，古代汉语和现代汉语难于截然区分，因为古汉语里的一些成分在现代汉语仍然使用。这就更加增加了问题的复杂性。比如：怪，原来认为义项5（妖怪；怪异的人或事物）是非成词语素，未标注词类，现在认为是成词语素，增标了“名”（鬼怪、京城一怪）。这一改变是符合现代汉语的实际情况的。“怪”的名词用法，虽然上古汉语就已经存在（例如出自《论语》的“子不语怪、力、乱、神”），但是现代汉语直接继承了这一用法，在有些场合仍作为名词使用。一些多字条目，到底是词还是词组，也颇难定夺。例如：满堂彩、不得了、不敢当、不过意、不是味儿、不像话、不要紧、满负荷、双肩挑、悄没声儿、一锅煮、

一勺烩、第一把手、胳膊腕子、胳膊肘子、死乞白赖，这些多字条目，过去认为是短语，未标注词类；现在则承认它们是词，增标词类。既然承认上述这些条目是词，那么“转关系、坐月子、做手脚、做文章”似乎更应该承认是词，然而《现汉₆》与《现汉₅》都未标注词类。还有一些三字条目，如：核战争，弹簧门、弹簧秤、中国字，到底是词还是短语，也还可以进一步讨论。再如，“第一把手”，增标为名词，“第一夫人、第一时间”是不是也应增标为名词？所有这一切，都充分体现了汉语划分词与非词的困难。其次，是名、动、形三大实词的分界问题。汉语划分词类之难，很大程度上难在名、动、形的划分。名动之间、形动之间和名形之间都很难划分出清晰的界限。这种状况在本次词类标注变化的频次中体现得十分明显：《现汉₆》词类标注的变化虽然涉及9种词类，但名、动、形三类的变化频次，尤其是名词和动词的变化频次占绝对优势，其他词类，除副词外，变化频次几乎可以忽略不计。《现汉₆》词类标注的这些变化也有力地印证了沈家煊（2007，2009，2010a，2010b，2012）^②从不同角度所论证的一个重要观点：汉语里名词和动词属于“名动包含”模式，即汉语的动词是名词所包容的一个次类。

三、结语

一般认为，标注词类是一部词典的基本要素之一。Jackson（1985）^③指出：“一部词典可以不提供其他语法信息，但是标注词类是必不

① 吕叔湘：《汉语语法分析问题》，商务印书馆1979年版，第17页。

② 沈家煊：《汉语里的名词和动词》，《汉藏语学报》2007年第1期；沈家煊：《我只是接着向前跨了半步——再谈汉语的名词和动词》，《语言学论丛》2010年第40辑；沈家煊：《英汉否定词的分合和名动的分合》，《中国语文》2010年第5期；沈家煊：《名动词的反思：问题和对策》，《世界汉语教学》2012年第1期。

③ Jackson, Howard. Grammar in Dictionary. In Robert Ilson (ed.), *Dictionary, Lexicography and Language Learning*. Oxford: Oxford University Press, 1985, pp. 53-59.

可少的。”Beck (2002)^①也指出：“词类是词典编纂的要素。”词典标注词类，对于许多语言，如俄语、法语、德语、英语等而言，是比较容易的。然而，对于基本上没有形态变化的汉语来说，却是一项非常困难的工作。大概是因为这个原因，《现汉》最初（1960年“试印本”）只是对虚词标注词类，45年后，2005年出版的《现汉₅》才开始全面标注词类。七年后，《现汉₆》对《现汉₅》的词类标注予以修订。从这些修订我们能够更为真切地感受到汉语词类标注之难主要在于两条界线：一是名、动、形三大实词之间的界线，二是词与非词（语素和短语）之间的界线。虽然有此二难，凭藉《现汉》修订者的智慧和努力，经过本次修订，《现汉》的词类标注的准确性得到了较大提高。但是，《现汉₆》的词类标注仍存在需要进一步讨论的地

方。例如：“靡靡之音”标为形容词显然欠妥。“食用、手洗、双栖、速溶、自助”都标为动词，显然标为形容词（非谓形容词）更好些。“搁不住、搁得住”，应不应该标注词类？“恒定”标为动词，改标为形容词是否更好些？“宣教”标为名词，改标为动词是否更好些？

本文作者：北京外国语大学语言研究所副教授，中国社会科学院研究生院语言系2008届语言学博士
责任编辑：马光

① Beck. *The Typology of Parts of Speech Systems: The Markedness of Adjectives*. New York: Routledge, 2002, p. 3.

An Analysis of the Statistics of the Revision to the Part-of-Speech Tagging in *The Contemporary Chinese Dictionary* (6th Edition)

Yin Hongbo

Abstract: *The Contemporary Chinese Dictionary* (6th Edition) made some revisions to the part-of-speech tagging of *The Contemporary Chinese Dictionary* (5th Edition). By overall statistics, I find there are four kinds of revision; adjusting the orders of the different parts of speech, changing the original part-of-speech tagging, adding and deleting some part-of-speech tagging. These revisions involve nine kinds of part of speech, but the nouns, verbs and adjectives, especially the former two kinds are revised most frequently, which strongly verifies the argument Shen Jiakuan has been advocating from different angles in recent years: in Chinese the noun contains the verb, namely the verb is a subcategory of the noun. The changes of the part-of-speech tagging of *The Contemporary Chinese Dictionary* (6th Edition) make us deeply feel that the difficulty of the part-of-speech tagging in Chinese mainly lies with two boundaries: one boundary is amongst the three kinds of notional words, i. e., the noun, the verb and the adjective; the other boundary is between words and non-words (including morphemes and phrases). The accuracy of the part-of-speech tagging in *The Contemporary Chinese Dictionary* (6th Edition) has been greatly enhanced, though it can be achieved still further.

Keywords: *The Contemporary Chinese Dictionary* (6th Edition); part of speech tagging; statistics and analysis