**[中文分词工具jieba中的词性类型](https://www.cnblogs.com/adienhsuan/p/5674033.html)**

jieba为自然语言语言中常用工具包，jieba具有对分词的词性进行标注的功能，词性类别如下：

|  |  |  |
| --- | --- | --- |
| **Ag** | 形语素 | 形容词性语素。形容词代码为 a，语素代码ｇ前面置以A。 |
| **a** | 形容词 | 取英语形容词 adjective的第1个字母。 |
| **ad** | 副形词 | 直接作状语的形容词。形容词代码 a和副词代码d并在一起。 |
| **an** | 名形词 | 具有名词功能的形容词。形容词代码 a和名词代码n并在一起。 |
| **b** | 区别词 | 取汉字“别”的声母。 |
| **c** | 连词 | 取英语连词 conjunction的第1个字母。 |
| **dg** | 副语素 | 副词性语素。副词代码为 d，语素代码ｇ前面置以D。 |
| **d** | 副词 | 取 adverb的第2个字母，因其第1个字母已用于形容词。 |
| **e** | 叹词 | 取英语叹词 exclamation的第1个字母。 |
| **f** | 方位词 | 取汉字“方” |
| g | 语素 | 绝大多数语素都能作为合成词的“词根”，取汉字“根”的声母。 |
| h | 前接成分 | 取英语 head的第1个字母。 |
| i | 成语 | 取英语成语 idiom的第1个字母。 |
| j | 简称略语 | 取汉字“简”的声母。 |
| k | 后接成分 |  |
| **l** | 习用语 | 习用语尚未成为成语，有点“临时性”，取“临”的声母。 |
| **m** | 数词 | 取英语 numeral的第3个字母，n，u已有他用。 |
| **Ng** | 名语素 | 名词性语素。名词代码为 n，语素代码ｇ前面置以N。 |
| **n** | 名词 | 取英语名词 noun的第1个字母。 |
| **nr** | 人名 | 名词代码 n和“人(ren)”的声母并在一起。 |
| **ns** | 地名 | 名词代码 n和处所词代码s并在一起。 |
| **nt** | 机构团体 | “团”的声母为 t，名词代码n和t并在一起。 |
| **nz** | 其他专名 | “专”的声母的第 1个字母为z，名词代码n和z并在一起。 |
| **o** | 拟声词 | 取英语拟声词 onomatopoeia的第1个字母。 |
| **p** | 介词 | 取英语介词 prepositional的第1个字母。 |
| **q** | 量词 | 取英语 quantity的第1个字母。 |
| **r** | 代词 | 取英语代词 pronoun的第2个字母,因p已用于介词。 |
| **s** | 处所词 | 取英语 space的第1个字母。 |
| **tg** | 时语素 | 时间词性语素。时间词代码为 t,在语素的代码g前面置以T。 |
| **t** | 时间词 | 取英语 time的第1个字母。 |
| u | 助词 | 取英语助词 auxiliary |
| **vg** | 动语素 | 动词性语素。动词代码为 v。在语素的代码g前面置以V。 |
| v | 动词 | 取英语动词 verb的第一个字母。 |
| **vd** | 副动词 | 直接作状语的动词。动词和副词的代码并在一起。 |
| **vn** | 名动词 | 指具有名词功能的动词。动词和名词的代码并在一起。 |
| **w** | 标点符号 |  |
| x | 非语素字 | 非语素字只是一个符号，字母 x通常用于代表未知数、符号。 |
| **y** | 语气词 | 取汉字“语”的声母。 |
| **z** | 状态词 | 取汉字“状”的声母的前一个字母。 |
| **un** | 未知词 | 不可识别词及用户自定义词组。取英文Unkonwn首两个字母。**(非北大标准，CSW分词中定义)** |

来源：<https://www.cnblogs.com/adienhsuan/p/5674033.html>

## 结巴分词词性对照表（按词性英文首字母排序）

### 形容词(1个一类，4个二类)

a 形容词

ad 副形词

an 名形词

ag 形容词性语素

al 形容词性惯用语

### 区别词(1个一类，2个二类)

b 区别词

bl 区别词性惯用语

### 连词(1个一类，1个二类)

c 连词

cc 并列连词

### 副词(1个一类)

d 副词

### 叹词(1个一类)

e 叹词

### 方位词(1个一类)

f 方位词

### 前缀(1个一类)

h 前缀

### 后缀(1个一类)

k 后缀

### 数词(1个一类，1个二类)

m 数词

mq 数量词

### 名词 (1个一类，7个二类，5个三类)

名词分为以下子类：

n 名词

nr 人名

nr1 汉语姓氏

nr2 汉语名字

nrj 日语人名

nrf 音译人名

ns 地名

nsf 音译地名

nt 机构团体名

nz 其它专名

nl 名词性惯用语

ng 名词性语素

### 拟声词(1个一类)

o 拟声词

### 介词(1个一类，2个二类)

p 介词

pba 介词“把”

pbei 介词“被”

### 量词(1个一类，2个二类)

q 量词

qv 动量词

qt 时量词

### 代词(1个一类，4个二类，6个三类)

r 代词

rr 人称代词

rz 指示代词

rzt 时间指示代词

rzs 处所指示代词

rzv 谓词性指示代词

ry 疑问代词

ryt 时间疑问代词

rys 处所疑问代词

ryv 谓词性疑问代词

rg 代词性语素

### 处所词(1个一类)

s 处所词

### 时间词(1个一类，1个二类)

t 时间词

tg 时间词性语素

### 助词(1个一类，15个二类)

u 助词

uzhe 着

ule 了 喽

uguo 过

ude1 的 底

ude2 地

ude3 得

usuo 所

udeng 等 等等 云云

uyy 一样 一般 似的 般

udh 的话

uls 来讲 来说 而言 说来

uzhi 之

ulian 连 （“连小学生都会”）

### 动词(1个一类，9个二类)

v 动词

vd 副动词

vn 名动词

vshi 动词“是”

vyou 动词“有”

vf 趋向动词

vx 形式动词

vi 不及物动词（内动词）

vl 动词性惯用语

vg 动词性语素

### 标点符号(1个一类，16个二类)

w 标点符号

wkz 左括号，全角：（ 〔 ［ ｛ 《 【 〖 〈 半角：( [ { <

wky 右括号，全角：） 〕 ］ ｝ 》 】 〗 〉 半角： ) ] { >

wyz 左引号，全角：“ ‘ 『

wyy 右引号，全角：” ’ 』

wj 句号，全角：。

ww 问号，全角：？ 半角：?

wt 叹号，全角：！ 半角：!

wd 逗号，全角：， 半角：,

wf 分号，全角：； 半角： ;

wn 顿号，全角：、

wm 冒号，全角：： 半角： :

ws 省略号，全角：…… …

wp 破折号，全角：—— －－ ——－ 半角：--- ----

wb 百分号千分号，全角：％ ‰ 半角：%

wh 单位符号，全角：￥ ＄ ￡ ° ℃ 半角：$

### 字符串(1个一类，2个二类)

x 字符串

xx 非语素字

xu 网址URL

### 语气词(1个一类)

y 语气词(delete yg)

### 状态词(1个一类)

z 状态词