

**Table 1**Taxonomy of Chinese words used in developing **MSRSeg**.

LW	Lexical Word	教授, 朋友, 高兴, 吃饭
FT	Factoid word	
dat	Date	10 月 11 日
dur	Duration	20 多分钟
tim	Time	十二点三十分
per	Percent and fraction	60%, 1/8
mon	Money	25000 美元
NUMBER	Frequency, integer, decimal, ordinal, rate, etc.	三次, 三个, 12.2 亿, 第二个
MEASURE	Age, weight, length, area, capacity, speed, temperature, angle, etc.	78 岁, 700 公斤, 2 亿公里, 1.8 亿公顷, 78 亿立方米, 每秒 2.89 米, 9 摄氏度, 5 度
ema	E-mail	jfgao@microsoft.com
pho	Phone, fax, telex	62617711
www	WWW	www.microsoft.com
NE	Named Entity	
P	Person name	李俊生, 亚历山大
L	Location name	蒙特利尔, 圣海伦岛公园
O	Organization name	中央民族乐团, 毕加索博物馆