

Addendum to the Chinese Treebank Bracketing Guidelines (CTB7.0)

Nianwen Xue, Zixin Jiang

September 18, 2010

This addendum should be used in conjunction with the Chinese Treebank Bracketing Guidelines:

Nianwen Xue, Fei Xia, Shizhe Huang, and Anthony Kroch. “The Bracketing Guidelines for the Chinese Treebank (3.0)”. Institute for Research in Cognitive Science Technical Report No. IRCS-00-08, University of Pennsylvania.

1 New phrase types

FLR: Fillers.

The FLR label is added to account for chunks of text that do not fit well into the syntactic structure. Nodes labeled *FLR* dominate chunks of text that are redundant in a sentence.

```
(IP (NP-TPC (NP (PN 我们))
  (CP (WHNP-1 (-NONE- *OP*))
    (CP (IP (NP-SBJ (-NONE- *T*-1))
      (VP (VV 在)
        (FLR (DP (DT 那个))
          (NP-OBJ (NN 大陆))))
        (DEC 的)))
      (NP (NN 部分)))
    (PU , )
    (NP-SBJ (PN 大家))
    (VP (ADVP (AD 只是))
      (VP (VV 希望)
        (IP-OBJ (NP-SBJ (-NONE- *pro*))
          (VP (VV 能够))
```

(VP (VV 去)
(VP (VV 发展)))))))))

INC: Incomplete sentences.

INC is a top-level node that is used to label incomplete sentences. It is different from FRAG in that the latter is syntactically incomplete but semantically and functionally is complete.

(INC (CP-CND (ADVP (CS 如果))
(CP (IP (NP-SBJ (PN 我))
(VP (VV 要)
(VP (VV 做)
(NP-PN-OBJ (NR 无极))))
(SP 的话))))
(PU 。)))

DFL: disfluency.

The DFL label is used to label parts of a sentence that are repetitions or restarts that are typical in transcribed speech.

(CP-Q (IP (NP-TPC (ADJP (JJ 这样))
(QP (CD 一)
(CLP (M 个)))
(NP (NN 作战)))
(PU ,)
(FLR (DP (DT 这个)))
(NP-SBJ (DNP (NP (PN 他))
(DEG 的))
(DFL NP (NN 目的)))
(PU ,)
(NP (NN 目的)
(NN 目标)))
(VP (VC 是)
(NP-PRD (PN 什么))))
(SP 呢)
(PU ?)))

2 New Functional Tag

A new function tag '-NOM' is added for clauses that clearly have a clausal structure but occur in a nominal context:

(IP (NP-SBJ (NP-APP (NP-PN (NR 台北))
(NP (NN 局长))
(NP-PN (NR 陈德佩)))
(VP (LCP-TMP (IP (NP-SBJ (-NONE- *pro*))
(VP (NP-TMP (NT 上午))
(PP-LOC (P 在)
(NP (NN 立法院)))
(VP (VV 进行)
(NP-OBJ (NN 民事)
(NN 报告))))))
(LC 时))
(PP-PRP (P 对)
(NP (DNP (NP-PN (NR 台铁))
(DEG 的))
(IP-NOM (NP-SBJ (NN 处理))
(VP (VA 不当))))))
(PU ,)
(PP-DIR (P 向)
(NP (PN 大家)))
(VP (VV 道歉)))
(PU 。))