

## 前端 C 语言代码关键总结

1) 分词模块以 `emByte *g_pText` 为输入，将结果写入到 `g_pTextInfo`，主要修改 `Pos` 和 `Len` 字段。以“图灵机器人好”为例，说明 `g_pTextInfo` 的构造规则。

序号	词性 Pos	词长 Len	对应的文本	说明
0	48	4	图灵	“图灵”为 2 个字，所以要占 2 行。空行用来占位
1	0	0		
2	48	6	机器人	“机器人”为 3 个字，所以要占 3 行。空行用来占位
3	0	0		
4	0	0		
5	1	2	好	“好”为 1 个字，只占 1 行

2) 韵律分析模块以 `g_pTextInfo` 为输入，将结果写回到 `g_pTextInfo`，主要修改 `BorderType` 字段。手动修改 `BorderType` 的值，可以显著地改变音频的韵律感。

3) 注音模块以 `g_pText` 为输入，将结果写入到 `g_pTextPinYinCode`。`g_pText` 中保存的是汉字的 GBK 编码，`g_pTextPinYinCode` 中保存的是对应汉字的拼音编码。

4) 替换分词模块、韵律分析模块和注音模块遵循以上基本原则，持续更新中。。。

分词模块和韵律分析模块都修改了链表 `TextItem *g_pTextInfo`，其结构体声明为

```
struct TextItem
{
    emUInt8    Pos;           //词性
    emUInt8    Len;           //词长
    emUInt8    TextType;      //内容类型
    emUInt8    BorderType;    //边界类型：0 字，1 LW，2 PW，3 PPH
    emByte     nBiaoDian[2];  //本词后面跟的标点
};
```