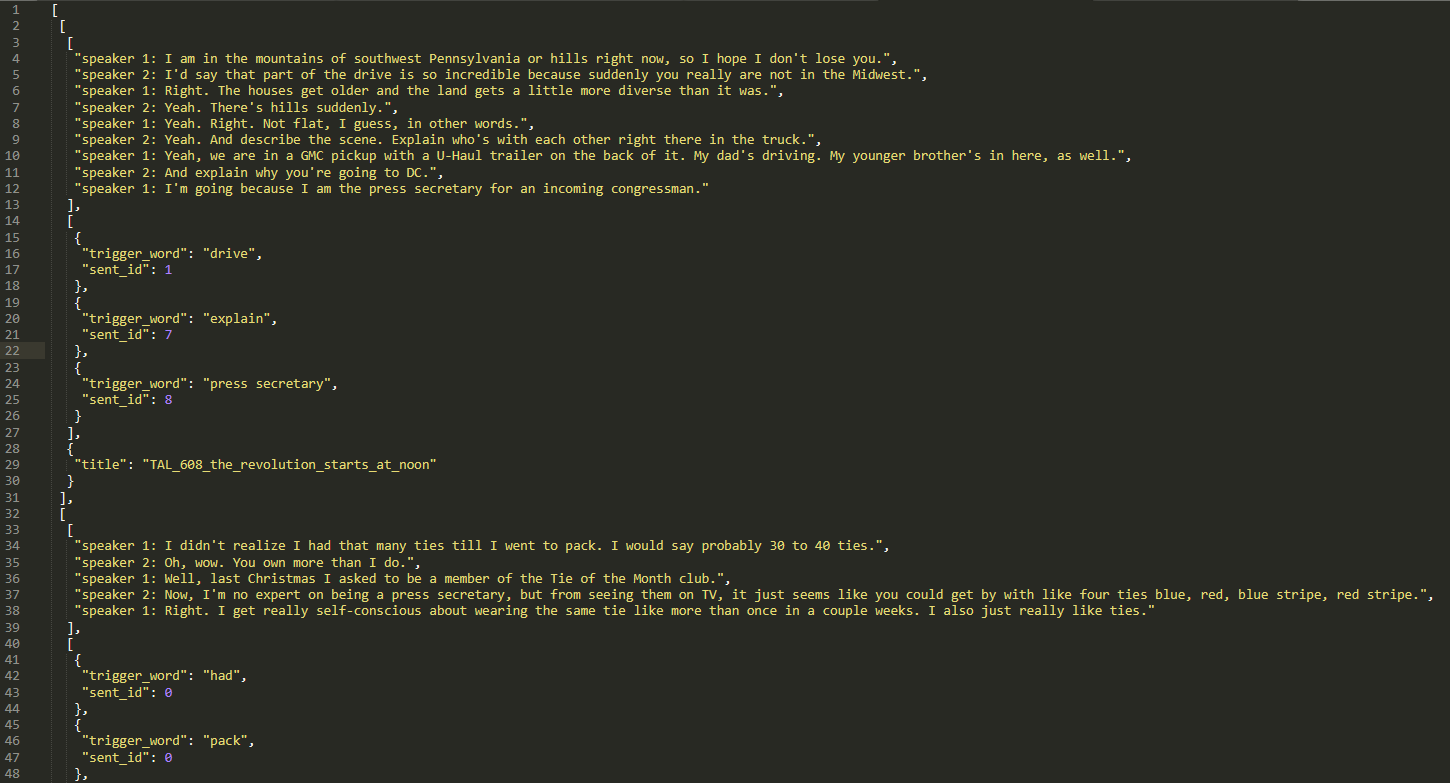
**Trigger\_word, 以句号为单位,能够代表整句话的动词，动词短语(主要是这前两个)，名词或者形容词。**

标注样例，见压缩包的sample.json, 一行代表一个speaker说的一个utterance



**注意事项！！！**

1. **格式一定和sample.json完全一致**，可以通过json.load(‘sample.json’)来检查格式是否正确。
2. 在一段对话当中，找到的trigger\_word(简单定义见文档最上方)的**同时记下sent\_id(或者说utterance id)，id从0开始算**。**感觉是trigger\_word的可以都标上去，后期会筛选**。
3. 在data/文件夹中有train.json, dev.json和test.json，从三个文件随机找**100个对话**进行标注。
4. 尽量找口水话比较少的对话，（尽量不找出现了较多诸如 how are you？ I am fine, thank you。 Hey， hello, ohhh… 口水话的对话。)
5. 标注的对话最好能超过十个utterance。
6. 一般一段对话有6-10个trigger\_word算是正常范围内。
7. 加油。