

如何使用CAT工具翻译字幕文件(之srt篇)

常见的字幕格式有srt，ass等。使用计算机辅助翻译工具(CAT)翻译字幕会大大提高效率，同时还能确保前后一致性，以Trados为例，如何翻译srt文件？

本文会提供几种思路。

一、使用Word对时间轴Markup

在Trados中默认是不处理Word的隐藏文字，根据这一特点，可以将srt文件用Word打开，然后在Word中进行处理。

test.srt 字幕样例：

1
00:00:05,049 -> 00:00:09,929
This is from Peking University.

2
00:00:09,964 -> 00:00:13,120
We have Trados ebook.

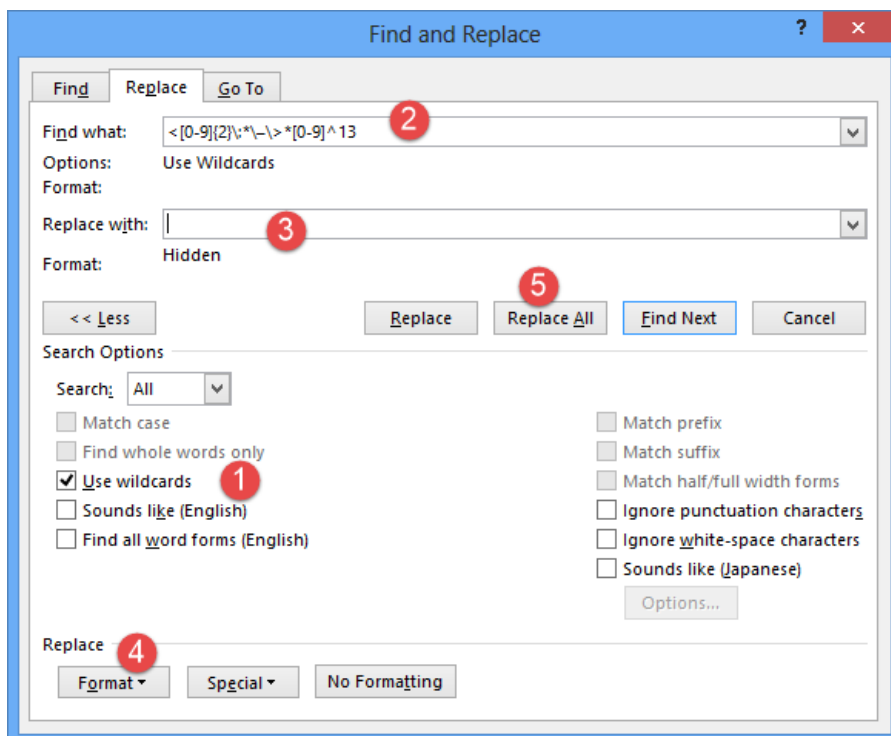
3
00:00:13,155 -> 00:00:17,113
Our eBook is for beginners.

不需要翻译的部分主要是句段号和时间轴

1
00:00:05,049 -> 00:00:09,929

因此可以在Word中使用查找和替换的方式，将所有的时间轴都找出，并将其格式设置为隐藏

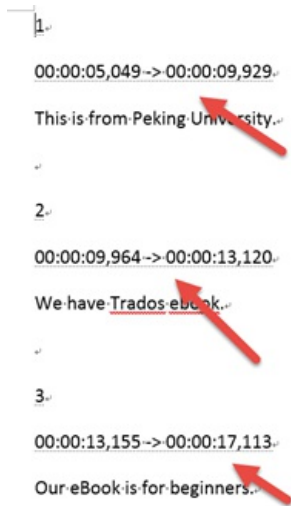
如下图所示：



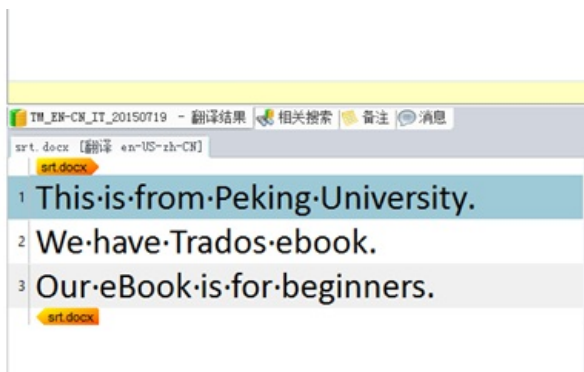
1. 勾选"Use wildcards"
2. 在"Find what:"中输入，<[0-9]*\-\>*[0-9]^13
3. 在"Replace with:"中，将格式设置为 Hidden
4. 点击"Replace All"

这样便可将所有时间轴选中，同时将其设置为隐藏格式，这样Trados Studio就不会将其抽取出来进行翻译了。

查找替换后的效果如下：



虚线表示隐藏文字，隐藏文字不会抽取出来翻译。在Trados中的效果如下：



翻译完成后导出译文，最后恢复隐藏文字便可以完成翻译。

完整步骤，可见如下视频。

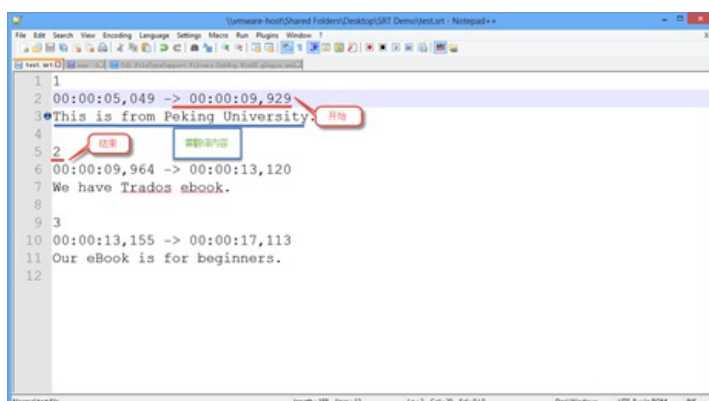
二、直接解析srt文件

基本思路，SDL Trados 可以指定开始和结束的位置，然后将其中需要翻译的内容抽取出来。一个简易的开始和结束规则如下：

开始: > 00:00:09,929

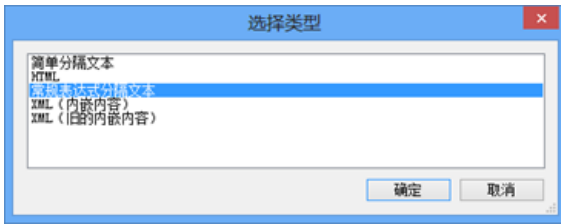
结束: > 2

在这个中间是需要抽取出来的翻译的。据此可以新建正则表达式匹配开始和结束的部分。

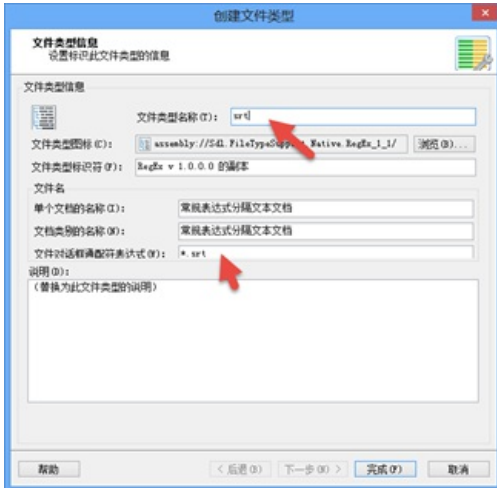


在Trados中的操作步骤：

1.



2.

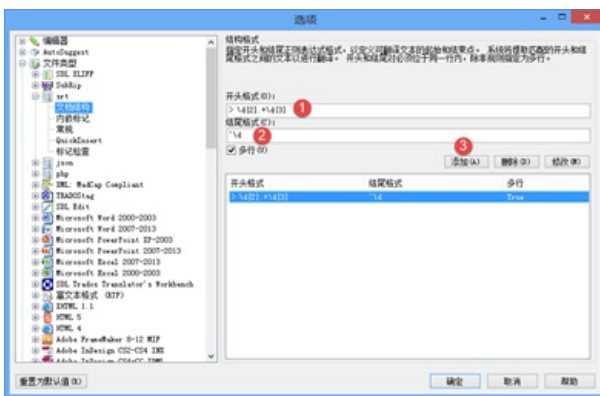


3.

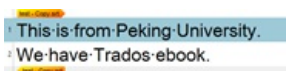
开始格式：> \d{2}.\d{3}

结尾格式：^\d

勾选多行。




4. 这时Trados 已经可以直接解析srt文件，解析效果如下：



细心的读者应该已经看出问题了，最后一句丢失了。因为目前最后一句是不符合抽取规则的，所以无法抽取。

这个方法不够完善，需要本地化工程师在最后一行手动写上一个数字，才能正常。如下图所示：

```
1 1
2 00:00:05,049 -> 00:00:09,929
3 This is from Peking University.
4
5 2
6 00:00:09,964 -> 00:00:13,120
7 We have Trados ebook.
8
9 3
10 00:00:13,155 -> 00:00:17,113
11 Our eBook is for beginners.
12
13 4
14
```



这样便可正常将所有的句段抽取出翻译了，翻译后再把最后添加的数字删除。这个方法较为简单便捷。

三、使用subrip插件，轻松搞定srt字幕翻译

这个插件可以抽取出所有的句段进行翻译，还可以修复断行的问题，总体不错，推荐大家使用。

官方下载地址：<http://www.translationzone.com/openexchange/app/subripfiletype-472.html?action=Download>

或者从我的网盘下载：<http://pan.baidu.com/s/1sjQK5rR>

四、SDL Trados 2015 或 Google Translate Toolkit

如果你使用的是SDL Trados 2015 或者是Google Translate Toolkit就不需要这么费心了，目前这两款已经可以直接翻译srt字幕了，省去不少烦恼，不过貌似也少学不少东西啊。

总结：

在使用工具的过程中经常发现这个现象，原先一些非常难以处理的文件格式，例如FrameMaker的mif文件，教学的时候常常需要讲2-3个小时才能说清楚，而随着工具的发展mif文件的翻译已经变得非常轻松。Srt字幕的翻译也是如此，早期在SDL Trados 2007的时候常常使用Word来做Markup，后来便不再依赖Word，而是Trados Studio中直接用正则表达式解析，再后来工具本身便直接支持了.....

工具的直接支持带来的便利性对最终用户来说是极好的，但是在教学的时候实际上并不太好，因为省去很多中间过程，学生反而不知道处理的常用方法了。例如，遇到ass格式的字幕，这时没有插件怎么办，或者同样是srt文件，如果贵司只有Alchemy Catalyst这时又改如何去处理呢？