

YouTube に翻訳字幕をつけるのは 結構大変でしたの話

おがさわらなるひこ

@ 小江戸らぐ 5月のオフな集まり (第 214 回)

LibreOffice 日本語チーム YouTube 公式チャンネル

<https://www.youtube.com/channel/UCJjOaLmTjYu4129vzZLmMNg>

100 人超えるとカスタム URL が作れるらしい

OSPN.jp の slack #random にて

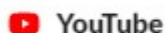


Jun Nogata 21:01

この流れなら言える（流れてない

LibreOffice日本語チームのYouTubeチャンネル登録してください。カスタムURLをつけたいです

<https://www.youtube.com/channel/UCJjOaLmTjYu4129vzZLmMNg>



YouTube

LibreOffice日本語チーム

LibreOffice日本語チームが主催したイベントなどの動画を公開します。



宮原さんが Twitter で宣伝してくれて
100 人突破！



宮原 徹 @tmiyahar

Apr 26

あと少しで100人チャンネル登録でカスタムURLが取得できるので、ご協力よろしくお願いします。

LibreOffice日本語チーム

[youtube.com/channel/UCJjOaLmTjYu4129vzZLmMNg](https://www.youtube.com/channel/UCJjOaLmTjYu4129vzZLmMNg) @YouTube
より

💬 4 🔄 10 ❤️ 14 ...

人数増えたのはいいがコンテンツ充実させたい

独自コンテンツは私以外の人に任せるとして
去年の AsiaCon(*) の基調講演、
二つともちょういい話だったので
あれに字幕つけるだけでも
コンテンツとして魅力的になるのでは……？ 🤖

↑ GW を使ってやりました。
宣伝エントリもブログに書いたよ ↓

<http://naruoga.hatenablog.com/entry/2020/05/04/215840>

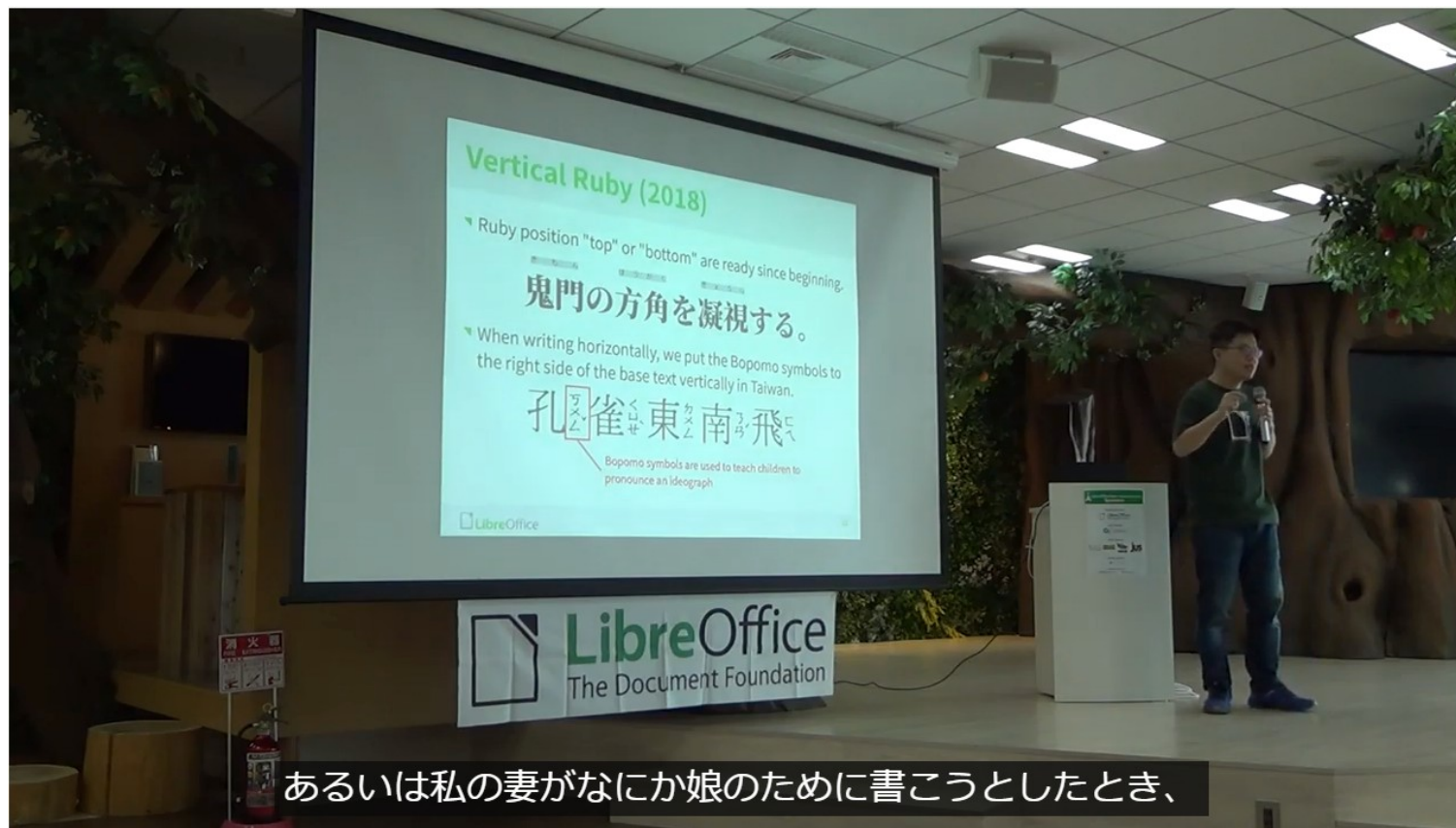
(*)LibreOffice Asia Conference 2019 Tokyo

<https://conf.libreoffice.jp/>

LibreOffice CJK Bugs, Fixes, and Stories

/ Mark Hung

<https://www.youtube.com/watch?v=FvOkLdTEguo>



LibreOffice, the many different faces of a global community / Italo Vignoli, Lothar Becker

<https://www.youtube.com/watch?v=QS-Zz-2ovo0>



ぶっちゃけ

……すごい大変でした……

2本合わせて 15 ～ 20 時間はかかってるかな？？

労力に見合う View 稼げてるかどうか謎

もしかして（もしかしくなくても）オレ無駄なことしてるのでは、感……

▼ 21/06/26
LibreOffice Asia Conference 2019: LibreO

63 回視聴・2019/06/26

なのでせめてここで愚痴をいう（のと宣伝）

やったこと

1. YouTube で自動音声認識で字幕付与
2. 音声を聞きながら正しいものに直す
3. ファイルに出力
4. 翻訳ファイルを作る
5. アップロード
6. 細かい修正

1. YouTube で自動音声認識で字幕付与

これは単に字幕付与 ON にして動画流すだけなので簡単

2. 音声を聞きながら正しいものに直す

YouTube のクリエイターツール（旧 UI）を利用

今は YouTube Studio らしいが、字幕編集は旧 UI じゃないとできない

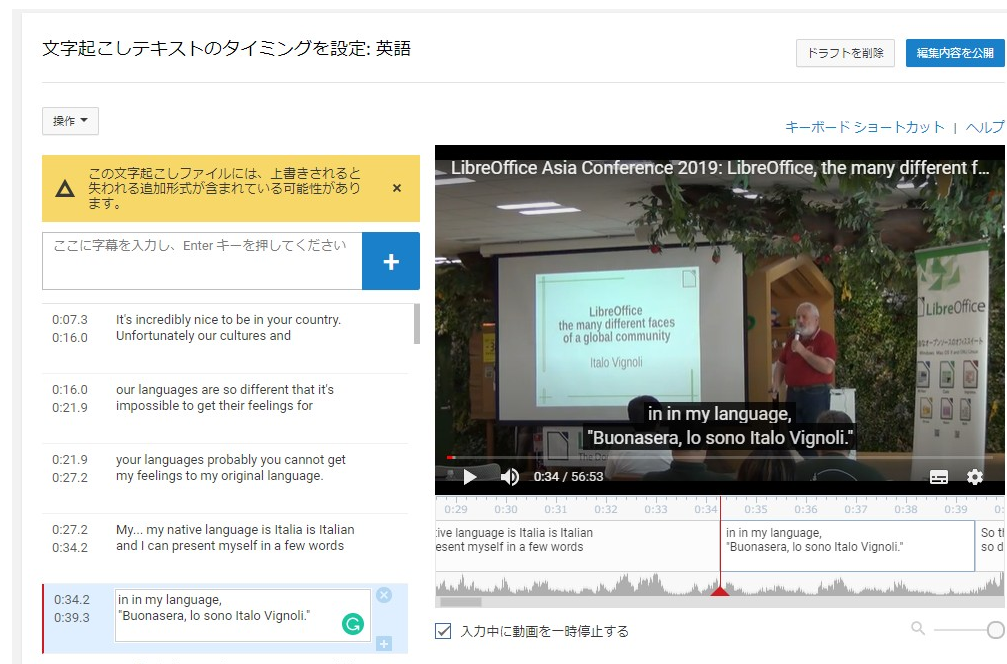
直すべきこと

句読点などの付与

字幕の切れ目

誤認識などなど

これらは別工程というより
なんとなく行きつ戻りつ



2. 音声を聞きながら正しいものに直す 句読点などの付与

全部小文字 & 句読点、疑問符、感嘆符など一切つかないので
とりあえずがつがつつける
切れ目修正と一緒にやっちゃうほうがいいかも

2. 音声を聞きながら正しいものに直す 字幕の切れ目

YouTube 先生の切り方は なんだか適当

意味があるところで切れるよう
に直す

だいたい2行ぐらいを目安

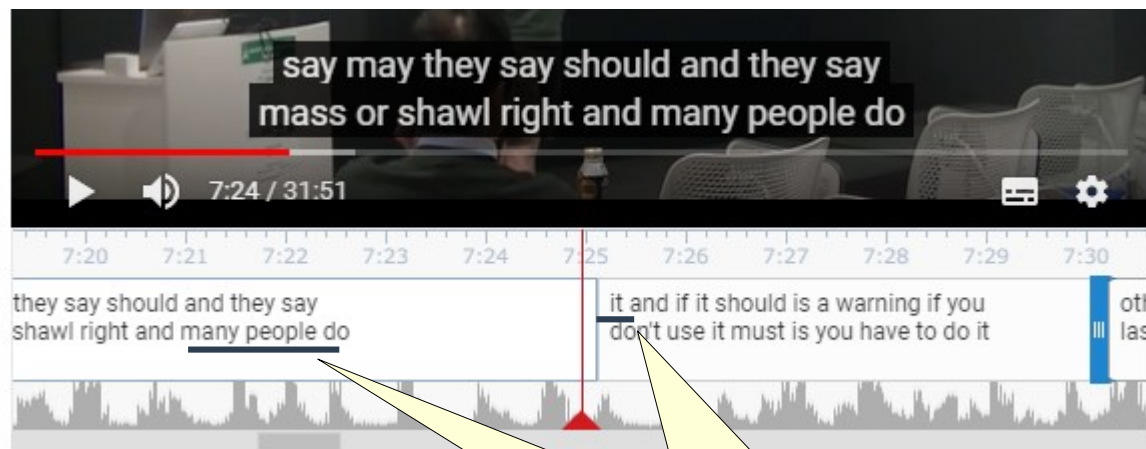
Enter キー押下で現在の再生位置でブロック分割

が、なんだけ変なところで切れることも……

よくわからん

文頭 Backspace 押下で前のブロックと結合

あとスライダー上 Drag & Drop で切る位置修正もできる



なんで
"many people do it"
の "it" だけ
泣き別れ……？

2. 音声を聞きながら正しいものに直す 誤認識などなど

肌感、認識率 8~9 割ぐらい？ 話し手に依存

英語のヒアリング頑張りながら拾いなおす

スロー再生を駆使（どうしてもわからないところは 0.5 倍速）

煮詰まったときにはお茶でもするとふと思いつくことがある

固有名詞

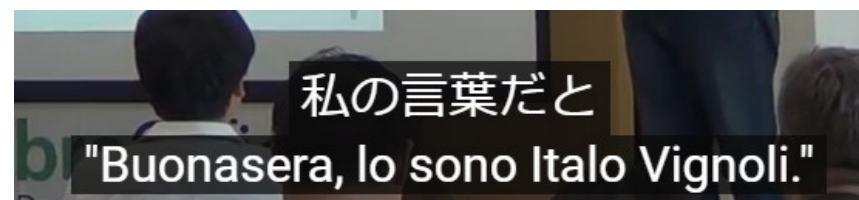
たとえば「LibreOffice」は全滅

人名はがんばって Google 検索して登場人物を探す

外国語

文脈からそれっぽい言葉を機械翻訳で検索して探す

スマホの Google 翻訳アプリを PC スピーカーに当てて聞き取り



たまに一部分ごっそり欠落してることもあるのでそこは自力で頑張れ

3. ファイルに出力

…… というか誤操作の Undo 動きが怪しいのでこまめに保存はしたほうがいいです

夏ぐらいに数時間作業した字幕を間違って削除し気力が尽きて停滞

これが公開に1年かかった理由……関係ないけど

出力されるファイル *.sbv は実際単なるテキスト

```
0:00:07.330,0:00:16.010  
It's incredibly nice to be in your country.  
Unfortunately our cultures and
```

```
0:00:16.010,0:00:21.920  
our languages are so different that it's  
impossible to get their feelings for
```

```
...
```

4. 字幕の翻訳

***.sbv ファイルを適当な名前のコピる**

あとは戸田奈津子になった気持ちで頑張る

機械翻訳使ってもいいけど話し言葉は正確じゃないので翻訳結果もヘンになることもある

そのまんま使うのはおススメできない

なるべく英語と語順合わせる

でも、ちょっと齟齬が出ても、字幕として一塊の意味が取れるようにするほうがよい

前の字幕と一緒に読んで意味を理解するほど人類は賢くない

5. アップロード

「言語の追加」で日本語作って

「ファイルをアップロード」して

「字幕ファイル」を選んで

作った *.sbv ファイルをアップロード

この工程は簡単！

…… 限定公開とかできんのかな？

よくわからないので目をつぶっていきなり publish しちゃった

6. 細かい修正

とりあえずぼーっと眺めておかしなところを直す

sbv ではいいと思っても字幕になってみると
微妙ってこともある

人に見てもらうのは効果的

オンラインエディタで直すより sbv 修正してアップロードし
たほうが無難

お悩みとかなんとか

言い間違いや言い淀みをどうする???

時勢間違えて言い直したり

uh... ah... so, so, so... など

原文書き起こしは（少し整理して）残した

全部の文頭に so 入ってるとちょっと考えちゃうけど……

俺の英語もそうだ！

翻訳だとそこらへんはバッサリカット

しゃべってるのに字幕進まないってこともでてしまうけど、

それはしょうがないかなーと

YouTube 先生の聞き取りの補助に AWS Transcribe 使ってみた

結果 JSON なのがちょっとね……

品質は YouTube とあんまり変わらなかったんで、次はないかなー

まとめ

LibreOffice 日本語チームのチャンネル登録 159 名わーい

みなさん協力ありがとー

LibreOffice Asia Conference 2019 の基調講演に字幕つけた

死ぬほど大変だったのでもうやらないというお気持ち

というかそれはわたくしの聞き取り能力の問題はありますね……

すみませんすみません

もしお仕事っぽい話だったらプロにやってもらうのがいいんじゃないかな？？？

次は日本語講演の英語字幕化にチャレンジしたい

こっちは聞き取りはずっとらく……な、はず！