

# Long video subtitling

Is there any difference?

Luca Braglia

July 4, 2021

# Introduction

## Setting

Let's say you want to create subtitles in your language for Gary Yourofsky's mix (3:37:42)<sup>a</sup>:

- how to do that?
- is the process different from short videos subtitling?

---

<sup>a</sup>Same for Workshop 2.0 (2:40:40) or Paul and Asa interview with David Ramms regarding rumors and accusations (3:02:35)

## Sections

- 1 quick background infos
- 2 how do we do it?
- 3 another possible way
- 4 subtitling nirvana

# Some background infos: what is subtitling btw?

## File formats

If you need:

- spreadsheet you create an `xlsx`;
- document go for `docx`;
- subtitles the most common format is `srt`.

## Subtitling in a nutshell

The creation of a proper (timing, translation) `srt`.

## How is it used?

Once you produced this file:

- you “pass it” to a player (eg VLC) and it display on the screen;
- in our case you send this file to the AV YouTube admins (Hudson) and they upload it in the main YT video.

## What is an srt?

It's a *text* file (eg you can edit with Notepad, but there are definitely better tools) with a strict *format*. An example from workshop 2.0 (starting seconds) follows:

1

00:00:00,000 --> 00:00:06,000

Mi chiamo Paul Bashir e sono il cofondatore  
e direttore di Anonymous for the Voiceless.

2

00:00:06,700 --> 00:00:12,300

Sono diventato un attivista per i diritti  
degli animali nel Giugno del 2015.

3

00:00:12,699 --> 00:00:23,199

Ho fondato AV con la mia compagna Asal, che ho incontrato  
poco prima che lanciassimo AV nell'Aprile del 2016.

# Workloads and feasibility

- producing a 5 minute length subtitle (good in translation quality/timing) is feasible *for a single person in a timely manner*;
- producing a 2-3 hours subtitle is definitely not;
- therefore a sensible solution could be to split long source subtitles in 5-min length chunks and distribute them across a team of translators/revisors (and this is what we're currently doing)

# Part I

## Our setup

# Overview: Team, skills and team creation

## Team and skills

Currently composed by three type of figures:

- 10+ translator: linguistic skills, basic computer skills (browser)
- 3 revisors: linguistic skills, intermediate computer skills (browser, editing srt with ad hoc software)
- 1 repository admin: some programming skills

## Team creation

- The repository admin and revisors were already available: translators were found asking for help to all Italian organizers (first wave) or directly using the “Resources for activists” Telegram channel (second wave).
- New aspiring translators are evaluated by our revisors and, if ok, admitted to the Telegram groups involving language localization (both for short and long videos).

# Overview: documentation

Once added to the groups all the participants have links to the relevant doc. An english translation is reported/linked below:

- for translators: [https://github.com/av-italia/subs/blob/main/doc/translate\\_en.md](https://github.com/av-italia/subs/blob/main/doc/translate_en.md)
- for revisors: [https://github.com/av-italia/subs/blob/main/doc/revise\\_en.md](https://github.com/av-italia/subs/blob/main/doc/revise_en.md)



# Overview: Software and services

We use:

- 1 Telegram (<https://telegram.org>) for communication;
- 2 GitHub for repository and subs translation (our repository is here <https://github.com/av-italia/subs>);
- 3 Aegisub (<https://aegisub.org><sup>1</sup>) for readability stats and final revision;
- 4 custom developed software for project management: in the past we have used an R package (<https://github.com/lbraglia/lbav2>), but from the next iteration we'll use its port (and generalization) to Python (<https://github.com/av-italia/pyavsubs>).

---

<sup>1</sup>Currently the site is under maintenance but the software can be downloaded here: <https://github.com/Aegisub/Aegisub/releases/>

# Our current workflow

## steps

- ➊ obtaining a starting point subs file (autogenerated english):
  - starting from scratch (no subs nor timing) would be probably infeasible in a timely manner;
  - several sources: YouTube, Downsubs etc.
- ➋ split it into chunks of 5 minutes (done by the software previously mentioned)
- ➌ translation
- ➍ linguistic revision
- ➎ readability revision
- ➏ final approval by our RO
- ➐ submission to AV YouTube admins

Let's have a quick view on the main ones ...

# 1. Obtaining a starting sub

## What I do:

- 1 download the video;
- 2 upload it to my YT channel (keeping it secret);
- 3 wait 2/3 days for YT to create automatic english subtitles;
- 4 download the subtitles (it's a little tricky here, more info at the end)

## Hint

Just in case, here you can find what we have used as a starting point<sup>a</sup>:

- Best of Gary Yourofsky — Speeches + Interviews:  
`https://raw.githubusercontent.com/av-italia/subs/main/source/gymix.srt`
- Every Accusation And Rumour About AV - LIVE With Paul And Asal: `https://raw.githubusercontent.com/av-italia/subs/main/source/arav.srt`

---

<sup>a</sup>Holding non vegans accountable 2 not reported because a) video has then been cut and timing would be wrong b) there is 2.5 now.

### 3. Chunk allocation

←  AV: Traduzioni e Sub YouTube  
20 membri, 2 in linea

Messaggio fissato  
Cmq approfitto ... @all, per questo giro: Qualora t...



14:16 ✓✓

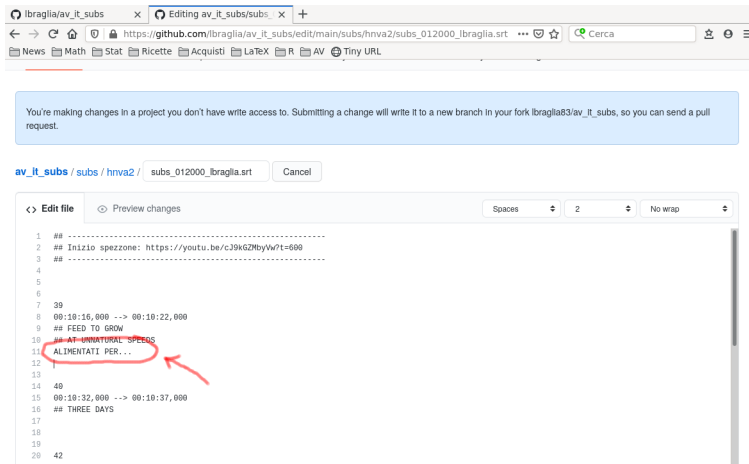
24 maggio

Agn25   
@lucailgarb #translate ↩ 3 10:57

Agn25   
@lucailgarb #translate  
subs/gymix/subs\_031500\_Agn25.srt 11:16 ✓✓

Agn25   
Luca Braglia  
subs/gymix/subs\_031500\_Agn25.srt  
completato 18:04

### 3. A day in the life of a translator



## 4. A day in the life of a revisor: phase 1

Spotify - La tua libreria x subs/subs/gymix at main x Update subs\_015000\_stu x Best of Gary Yourof

← → ↺ 🏠 🔒 https://github.com/av-italia/subs/commit/0b987642615d4a4b898b50d5a5c01cf80c17d581

📁 News 📁 Math 📁 Stat 📁 Science 📁 Ricette 📁 Acquisti 📁 LaTeX 📁 R 📁 Python 📁 AV 📁 casa 📁 finance



**AlessiaQ91** committed on 11 May Verified

1 parent 93e

Showing **1 changed file** with **46 additions** and **44 deletions**.

90 subs/gymix/subs\_015000\_sturiandre.srt

↑	@@ -8,7 +8,7 @@
8	8 01:50:10,640 --> 01:50:14,720
9	9 ## now i know these statements sound
10	10 ## pretty harsh and controversial to you
11	- Ora so che queste affermazioni ti suonano
11	+ Ora so che queste affermazioni vi suonano
12	12 piuttosto dure e controverse
13	13
14	14
↓	@@ -38,8 +38,8 @@ la compassione, dimostra come completamente irrazionali
38	38 01:50:30,480 --> 01:50:38,320
39	39

## 4. A day in the life of a revisor: phase 2

File Modifica Sottotitoli Timing Video Audio Automazione Visualizza Aiuto

What is animal rights activism

- Activism is speaking up for victims in the same way that you would want to be spoken for if you were in their position.
- It is being a medium of truth and enacting change by prioritising the victims and not the oppressors.

bene, quindi cos'è l'attivismo per i diritti degli animali?  
attivismo vuol dire parlare a nome delle vittime

2:00:18 19 20 21 22 23 24 25

Commento Default Edit Attore Eff 47

0 2:00:19.92 2:00:23.76 0:00:03.84 0 0 0

bene, quindi cos'è l'attivismo per i diritti degli animali? Nattivismo vuol dire parlare a nome delle vittime

2:00:19.933 - 216598 +13ms; -3827ms 100%

#	Inizio	Fine	CPS	Stile	Testo
1	2:00:04.40	2:00:11.84	9	Default	è la responsabilità per quello che l'Netanno facendo agli animali. Andiamo avanti.
2	2:00:19.92	2:00:23.76	23	Default	bene, quindi cos'è l'attivismo per i diritti degli animali? Nattivismo vuol dire parlare a nome delle vittime
3	2:00:23.76	2:00:27.36	23	Default	allo stesso modo in cui si vorrebbe che qualcuno l'Nparlasse per noi se fossimo nella loro posizione
4	2:00:28.48	2:00:35.20	13	Default	è essere mediatori di verità e innescare il cambiamento l'Ndando priorità alle vittime e non agli oppressori
5	2:00:36.00	2:00:42.88	9	Default	avete un minuto a partire da adesso l'NNon ho niente da aggiungere nemmeno qui
6	2:00:51.52	2:00:55.28	16	Default	Potrei passare a 30 secondi perché non voglio l'Ntrascinarla troppo a lungo
7	2:00:55.28	2:01:01.68	11	Default	il l'Ndeo, voglio essere il più conciso possibile l'Nma voglio essere sicuro di coprire ogni
8	2:01:01.68	2:01:06.72	10	Default	do l'Nda o questioni che potrebbero venire fuori. Voglio l'Nessere il più accurato possibile nel chiarire questo.
9	2:01:06.72	2:01:21.84	5	Default	Penso che sia il caso di ridurre il tempo a disposizione l'Nper voi ragazzi a 30 secondi invece che un minuto.
10	2:01:31.76	2:01:37.60	9	Default	Bene, andiamo avanti. Si ridurrà il tempo a l'N30 secondi tra ogni slide
11	2:01:37.60	2:01:42.56	15	Default	per permettervi di intervenire se volete l'Nche discussa di qualcosa più specificamente
12	2:01:45.52	2:01:49.84	8	Default	Carlo dice che è tutto chiarissimo. perfetto

# What do I do behind the scene?

```
l0m740n:~$ cd av_it_subs
l0m740n:~/av_it_subs$ inv gymix
=====
MAIN MENU (PRJ = GYMIX)
=====

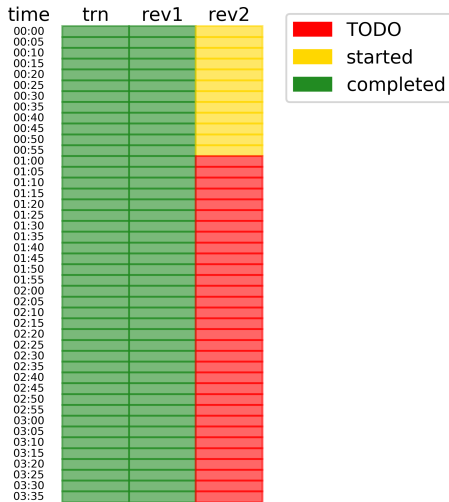
1. Setup
2. Create sandboxes
3. Assign TRN or REV2
4. Mark progresses
5. Monitoring
6. List available REV1
7. List available REV2
8. Make final srt
9. Final SRT stats
10. List assignee
11. List users

Selection (0 to exit): █
```



# Monitoring

gymix: avanzamento al 01/07/2021



## Part II

Another way

# Another possible way

## Autotranslated subs

Sometimes YT give the autotranslated subs (when english automatic subtitles are available or at least one sub/language is uploaded).

## An idea (not tested)

Use autotranslated subtitles in your language as a starting point, where available.

## Auspicious pros

- editing should be feasible, or at least it should be definitely less burden compared to full translation (hopefully);
- maybe (stressed) there's no need to split the source and arrange a team.

# How to obtain autotranslated subs in your language

You can download automatic translated subs using Downsubs (at the moment only the first and third link work):

- Holding non vegan accountable 2:  
`https://downsub.com/?url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DcJ9kGZMbyVw`
- Best of Gary Yourofsky: `https://downsub.com/?url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3D3gLOVtbG6QI`
- Every Accusation And Rumour About AV:  
`https://downsub.com/?url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DvQe-AGDKyzU`

## Part III

# Subtitling Nirvana

# Some thoughts for tomorrow ...

What would be the best possible solution/setup (IMHO)?

- the “major” /most important videos are subbed in English (eg even for deaf people)
- we all (national translation team) start from the same validated content (timing/language) and translate it to our language

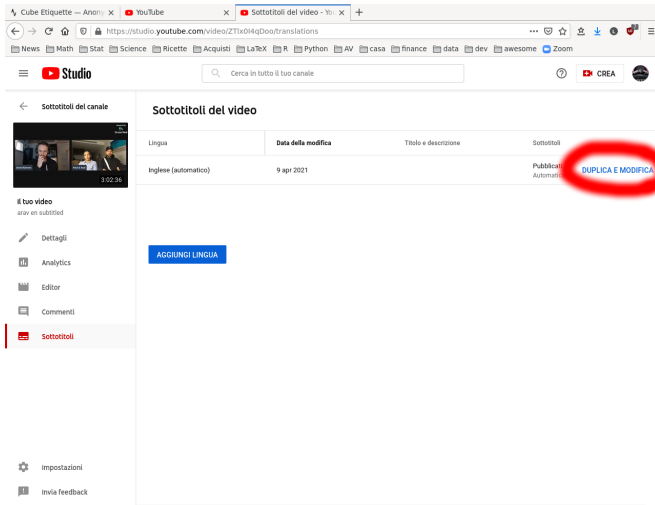
And that's all, thank you!

## Part IV

Misc



# Obtaining starting subs from YouTube (1/4)



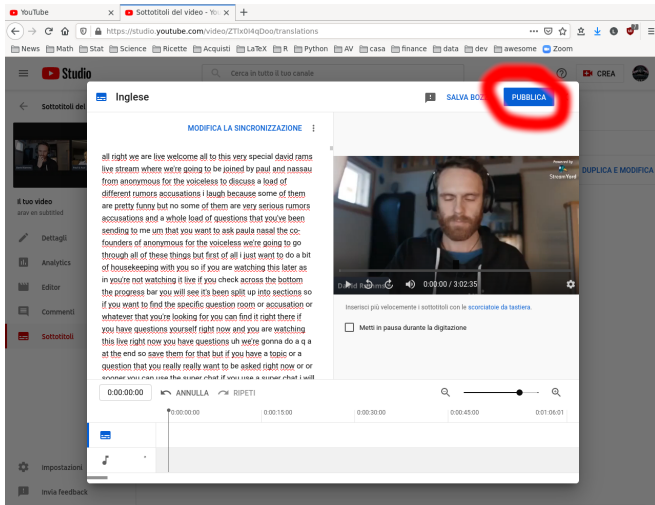
YouTube Studio interface showing the 'Sottotitoli del video' (Video Subtitles) page. The page displays a table with columns: Lingua, Data della modifica, Titolo e descrizione, and Sottotitoli. A row is visible for 'Inglese (automatico)' with a modification date of '9 apr 2021'. The 'Sottotitoli' column for this row shows 'Pubblicato Automaticamente' and a red button labeled 'DUPLICA E MODIFICA'.

Left sidebar menu items:

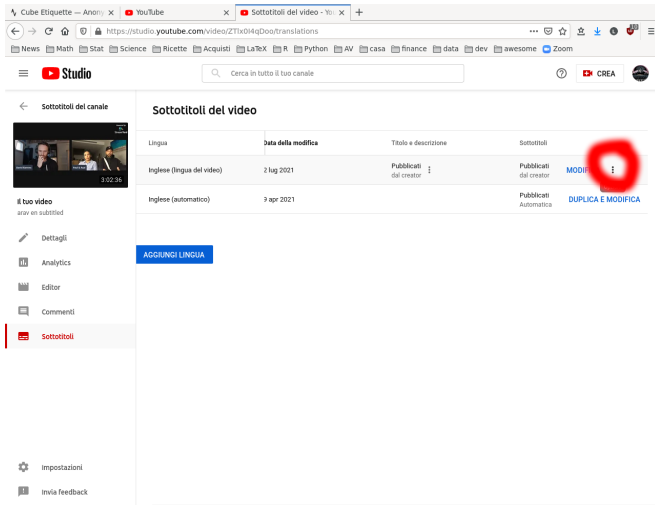
- ← Sottotitoli del canale
- il tuo video
- arav en subtitled
- Dettagli
- Analytics
- Editor
- Commenti
- Sottotitoli**
- Impostazioni
- Invia feedback

Bottom button: AGGIUNGI LINGUA

# Obtaining starting subs from YouTube (2/4)



# Obtaining starting subs from YouTube (3/4)



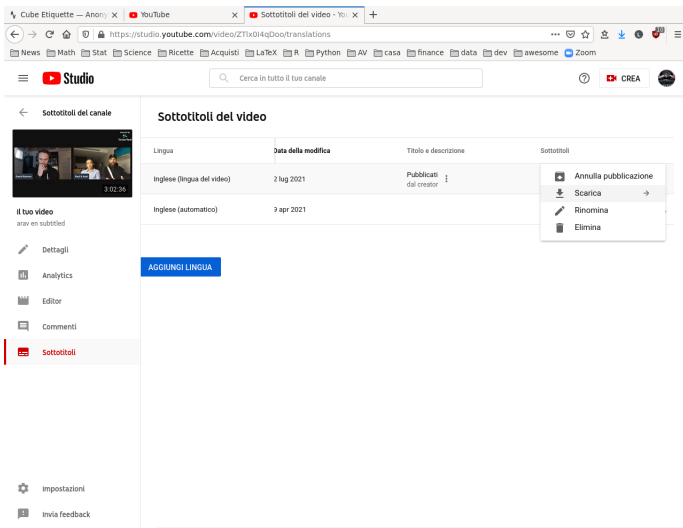
The screenshot shows the YouTube Studio interface for a video titled 'Sottotitoli del video'. The left sidebar contains navigation options: 'Sottotitoli del canale', 'il tuo video' (with a video thumbnail and 'arav en subtitléd'), 'Dettagli', 'Analytics', 'Editor', 'Commenti', 'Sottotitoli' (highlighted with a red vertical bar), 'Impostazioni', and 'Invia feedback'.

The main content area is titled 'Sottotitoli del video' and contains a table with the following data:

Lingua	Data della modifica	Titolo e descrizione	Sottotitoli
Inglese (lingua del video)	2 lug 2021	Pubblicati dal creator	Pubblicati dal creator <a href="#">MODIFICA</a>
Inglese (automatico)	9 apr 2021		Pubblicati Automatica <a href="#">DUPLICA E MODIFICA</a>

Below the table is a blue button labeled 'AGGIUNGI LINGUA'.

# Obtaining starting subs from YouTube (4/4)



YouTube Studio interface showing the 'Sottotitoli del video' (Video Subtitles) page. The page displays a table of subtitles for a video titled 'Sottotitoli del video'.

Lingua	Data della modifica	Titolo e descrizione	Sottotitoli
Inglese (lingua del video)	2 lug 2021	Publicati dal creator	<ul style="list-style-type: none"><li>Annulla pubblicazione</li><li>Scarica</li><li>Rinomina</li><li>Elimina</li></ul>
Inglese (automatico)	3 apr 2021		

AGGIUNGI LINGUA