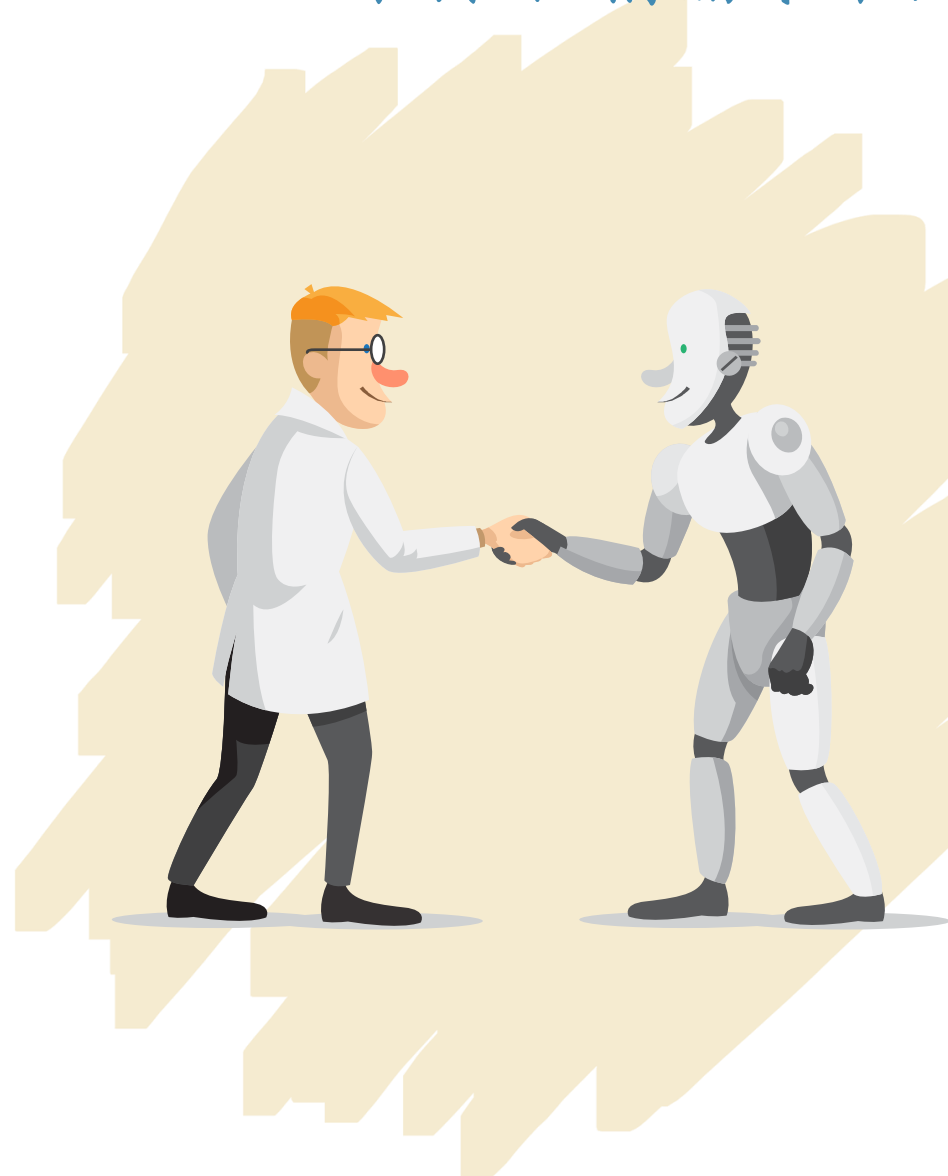


第12讲 影视字幕和 配音翻译技术

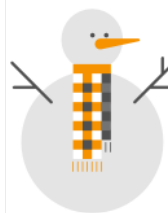
黄婕

2025-12-9 (4班)

2025-12-12 (3班)



本节 内容



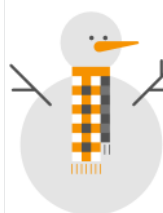
1. 视听翻译的内涵和范围

2. 字幕的特征和翻译要求



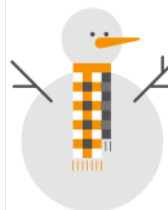
3. 配音的特点和翻译要求

4. 字幕翻译制作工具



5. 经典字幕翻译流程案例-
Aegisub

6. AI+字幕翻译流程



资源推荐



1. 视听翻译的内涵和范围

Question

1. 你觉得视频翻译和普通文本翻译有什么区别？

2. 你知道哪些专门的视频翻译工具吗？

3. 随着近几年音视频技术、人工智能的进步，你觉得视频翻译过程会有怎样的变化？

翻译对象的历史发展



软件本地化

网站本地化

视频本地化

1990年

2000年

2010年

Sample: <We bare bears> cartoon theme song



视频配音在不同地区的差异？

以下是美国动画片《We bare bears》的大陆、台湾和粤语三个版本的配音。

《We Bare Bears》讲述了三只拟人化的熊兄弟——灰熊Grizzly（大大）、熊猫Panda（胖达）和北极熊Ice Bear（白熊）在美国旧金山湾区生活的日常故事。它们试图融入人类社会，但常常因为熊的身份闹出各种搞笑又温馨的乌龙事件。

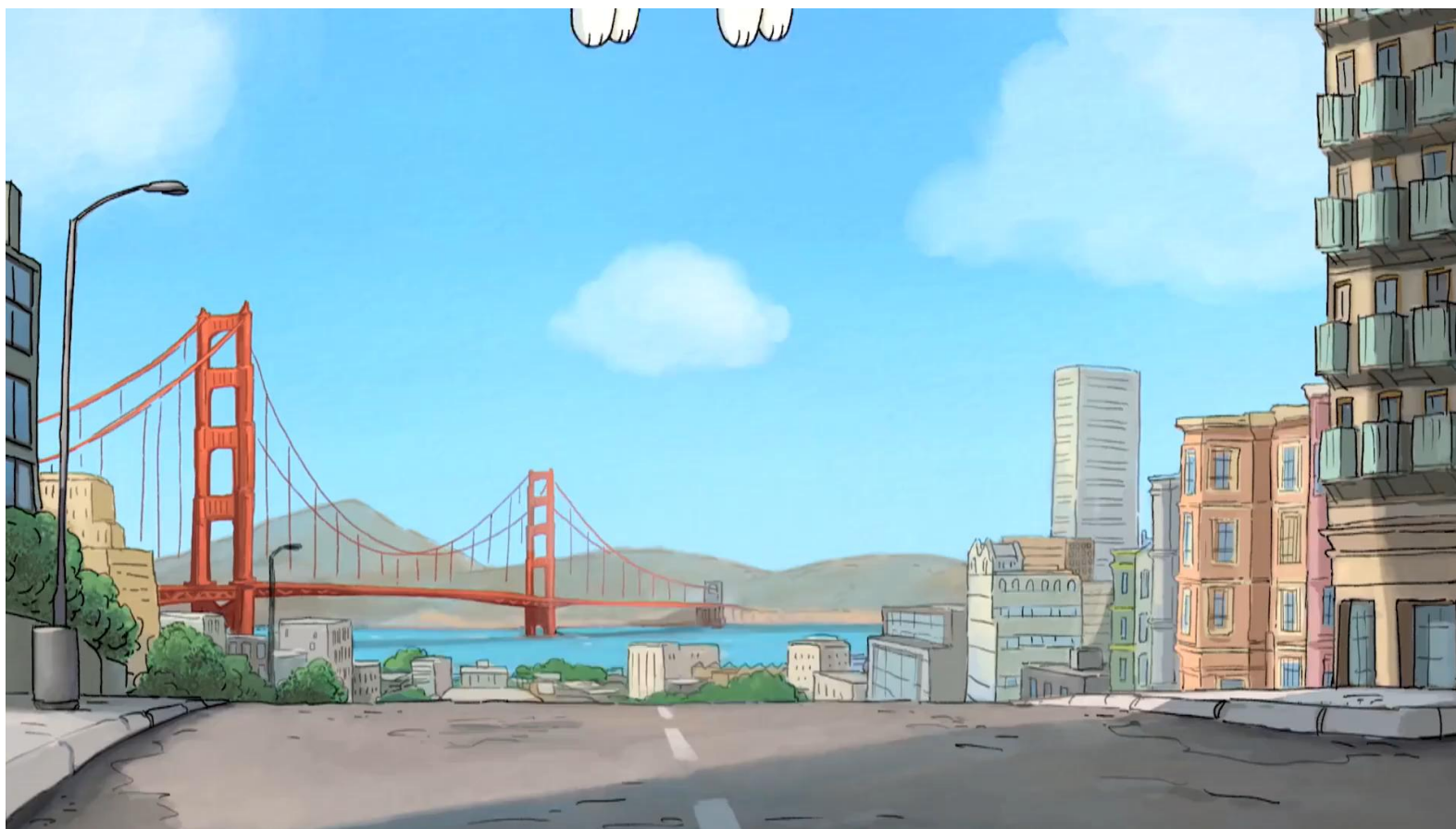
请留意三个版本的以下方面，找出差异：

1. 字幕文本的长度、风格、用词

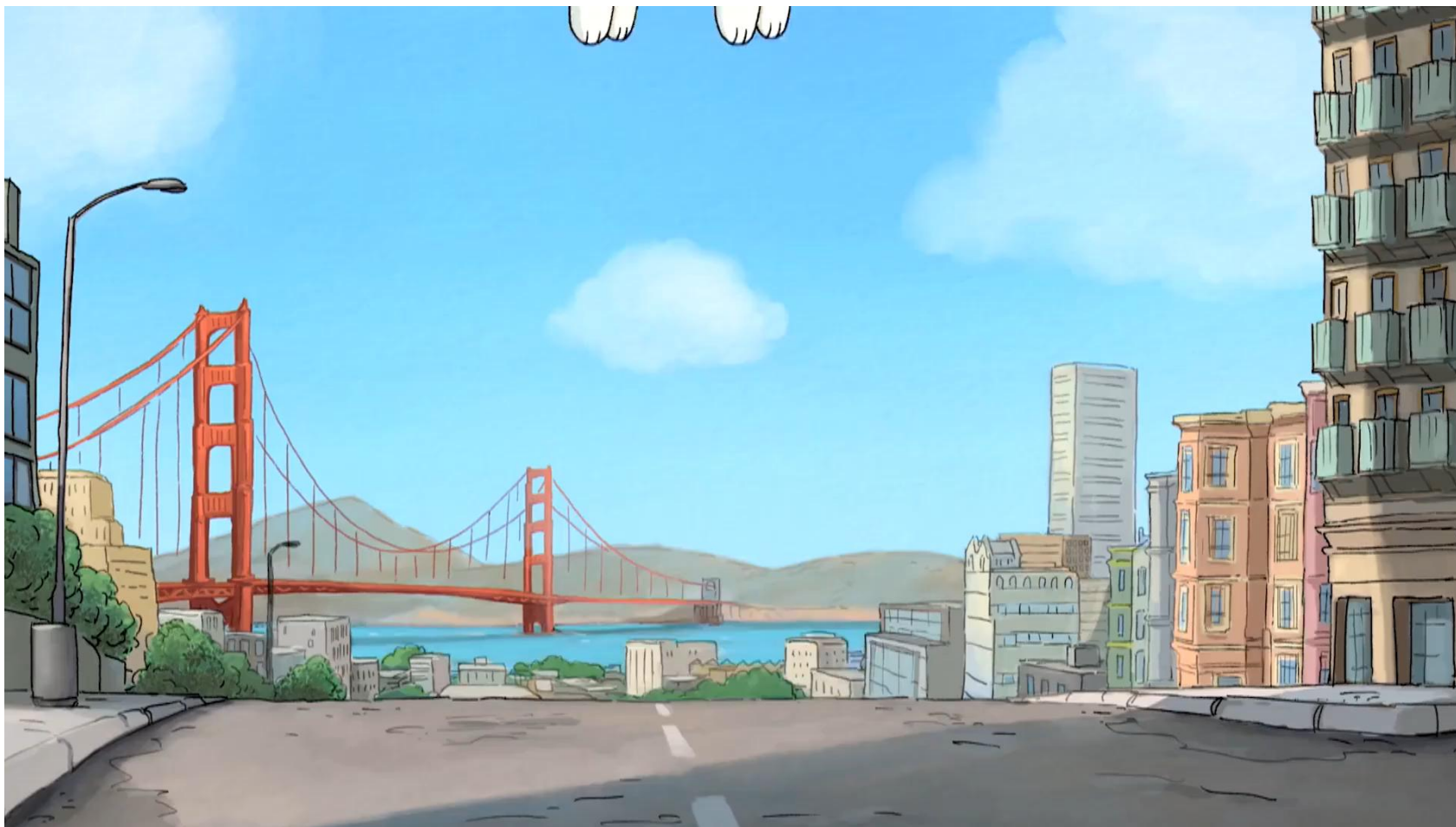
2. 配音的音色、情绪

3. 音乐的适配和调整

1. 大陆配音



2. 台湾配音



3. 粤语配音



Summary

字幕：

- 大陆：直译+口语化处理（来吧，咱都在！）
- 台湾：意译+年轻化网络用语（来撒野，熊熊遇见你）
- 粤语：半创作+方言特色（能令你欢心畅快任我唱，开朗：“朗”发/loŋ⁵/）

配音：

- 大陆：清亮的青年声线，强调“欢乐”集体感
- 台湾：接近台湾偶像剧夸张演绎
- 粤语：TVB配音传统，自然的说话式唱法，带懒音

音乐：

- 大陆：音调整体调高一度，更明亮
- 台湾：台词字数多，加速演唱
- 粤语：配乐增加和弦，因粤语常受港乐蓝调影响

视听翻译的内涵



多模态性:

- 包括语言、声音、图像等多种模态, 综合考虑以传达原始材料的意义和情感。

文化适应性:

- 翻译过程中需要考虑目标文化的特点, 对某些文化特定的元素进行适当的调整, 以确保目标观众能够理解和接受。

时效性:

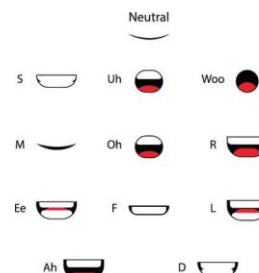
- 视听材料往往需要在特定时间内完成翻译, 以满足发布或播出的需求。

同步性:

- 字幕翻译需要与视频内容同步, 确保观众能够跟随剧情发展。

观众导向:

- 翻译应考虑目标观众的接受度和偏好, 进行适当的调整。



视听翻译的范围



视听翻译的特征和限制因素

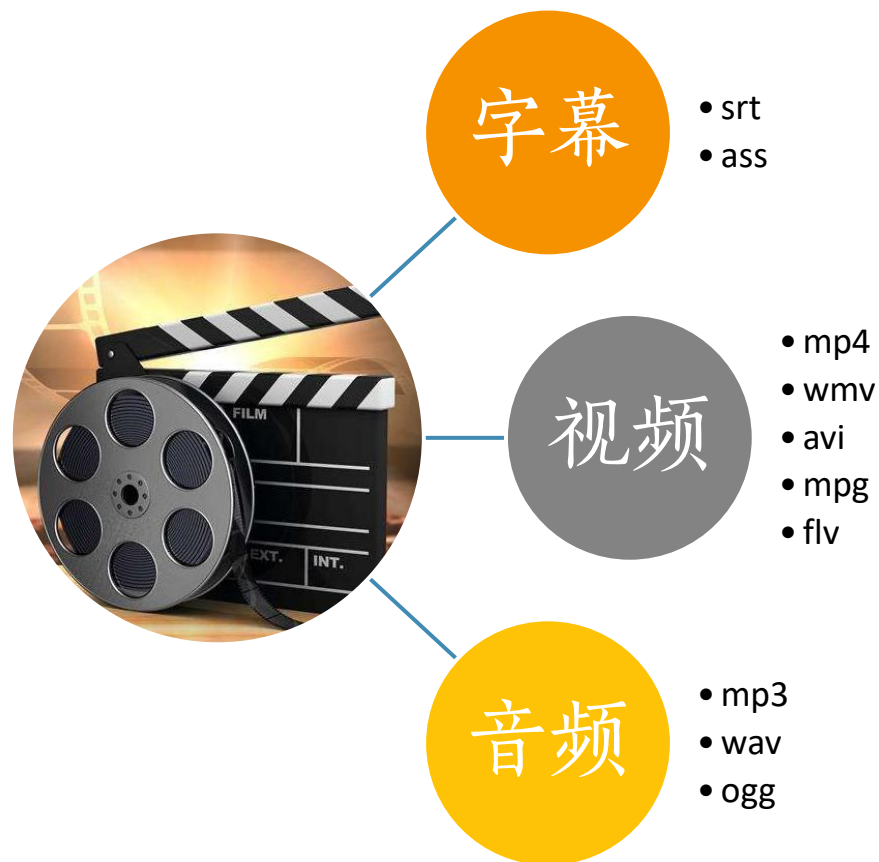
文本特征

- 即时性
- 大众性
- 口语化

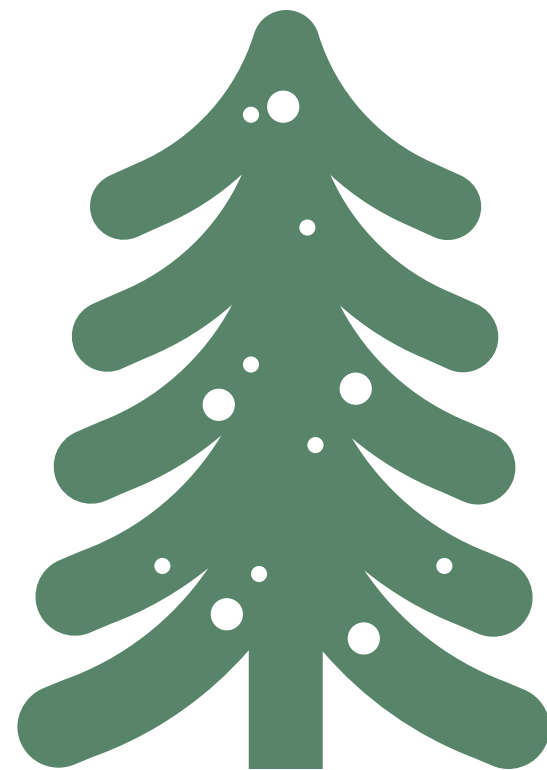
影响因素

- 时间
- 空间
- 文化

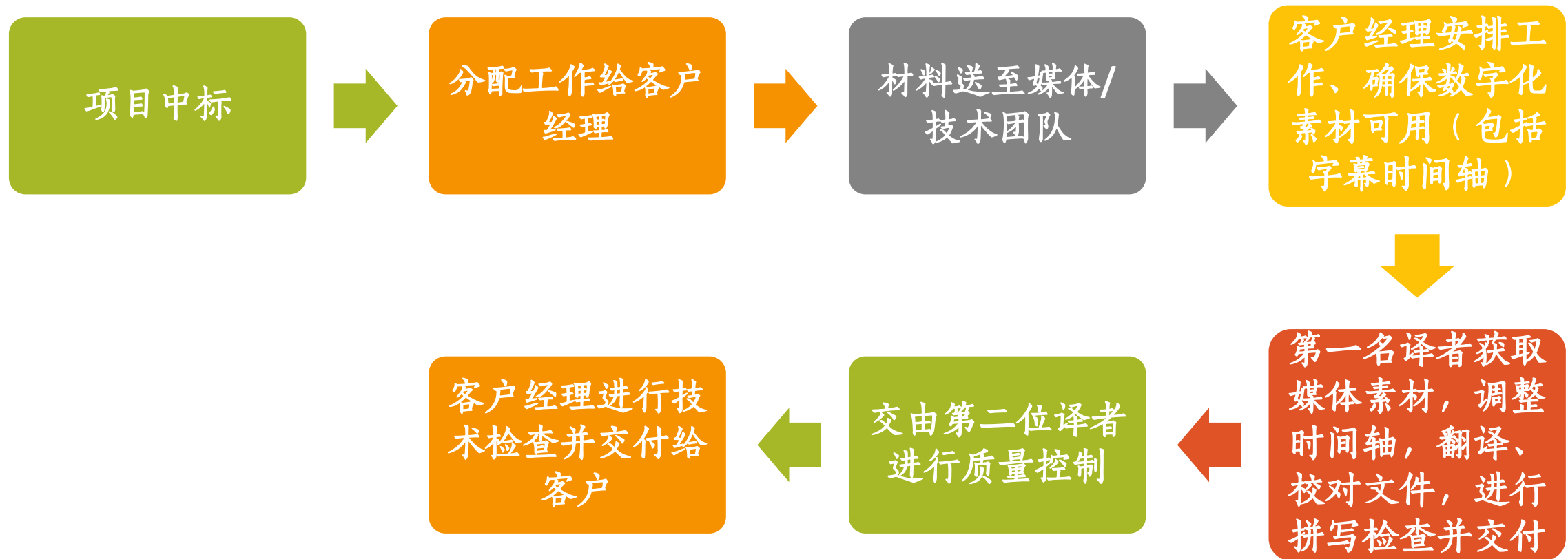
视听翻译的文件格式



2.字幕的特征和翻译 要求



字幕翻译流程



图表翻译自 **Figure 2.1 Typical subtitling workflow** from Cintas, Jorge Díaz, and Aline Remael. 2020.
Subtitling: Concepts and Practices. London & New York: Routledge, pp 38.

字幕文件的命名与位置

- 字幕文件的常见格式：SRT, ASS, SSA
- 字幕文件要与视频文件同名，并且放在相同文件夹下
- 使用合适的视频播放软件显示字幕（暴风影音等）
- 字幕文件的内容：时间码（开始，结束），字幕文字
- 字幕文件应完整，包括片头字幕、台词字幕、片尾字幕、说明性字幕和歌词或歌词大意字幕等，并包含准确的时码信息。硬字幕 vs 软字幕 有什么区别？

Burned-in subtitles
Open captions

Overlay subtitles
Closed captions



记得选择支持字幕
文件的视频播放器！

字幕翻译的要求-1

基本要求

- 字幕制作应对照原节目视频画面完成，应拼写准确、语法正确、断句合理，与视频画面以及配音保持同步，呈现效果风格应与原节目一致。

制作范围

- 对白、独白、旁白、解说、人物同期声等台词，应制作字幕，群杂声可根据需要制作字幕；
- 片名、片头字幕和片尾字幕应进行译制；
- 歌曲（含片头曲、片尾曲、插曲）名称、歌词或歌词大意，应制作字幕；
- 无字幕介绍的重要角色、节目背景、视频画面中清晰可见的文字（如手机短信、社交媒体信息、邮件、信件等），应制作字幕；
- 节目中出现的中国朝代、历史人物、历史典故等，应制作说明性字幕。

参考：广播电视和网络视听节目对外译制规范，国家广播电视总局

字幕翻译的要求-2

同步要求

- 字幕时间轴（记录字幕出现和消失时间的信息序列）应与视频画面、配音等保持同步。
- 字幕出现误差应不大于三帧，字幕消失误差应不大于三帧。

字幕呈现

- 字幕应优先保障信息的充分有效传达、受众清晰观看和充分阅读，应易于辨认、理解
- 位置：台词字幕应位于视频画面底部，在字幕安全框之内。
- 行数：台词字幕每屏一行，最多不超过两行。
- 字符数：字幕每行最多字符数可参考中文原字幕最大宽度。
- 完整性：每屏、每行字幕的分布应兼顾画面与句型、单词和语义的完整性。
- 进出屏幕方式：字幕应按照目标语言书写方向进出屏幕。
- 时间间隔：台词字幕时间间隔宜参照原节目，应与配音保持一致。

字幕翻译的要求-3

字幕呈现（续上页）

- 风格：字幕的效果风格等，应与原节目一致。
- 字体：宜选用清晰醒目的字体，不使用下划线。
- 字号：应兼顾目标语言阅读习惯与画面美观要求，不能遮挡画面重要信息，可参照原节目字幕字号大小。
- 颜色：台词字幕颜色宜为白色（容差范围为90%~100%），4K超高清HDR台词字幕亮度不应超过300cd/m²。字幕底色不宜使用彩色背景。
- 区分度：同一屏幕画面中出现的台词字幕和说明性字幕，应进行区分。两人对白字幕同时出现在同一屏幕画面中，应进行区分。
- 大小写：可根据目标语言习惯使用大小写。

例：BBC subtitle guidelines

<https://www.bbc.co.uk/accessibility/forproducts/guides/subtitles/>

BBC Sign in Home News Sport Earth Reel Worklife Travel ... 🔍

Accessibility for Products

Home Guidelines Champions Research Policy

OVERVIEW ▼

PRESENTATION ▲

- 2 Editing text ▼
- 3 Line breaks ▼
- 4 Timing ▼
- 5 Synchronisation ▼
- 6 Matching shots ▼
- 7 Identifying speakers ▼

Subtitle Guidelines 🔗

Version 1.2.2
January 2024

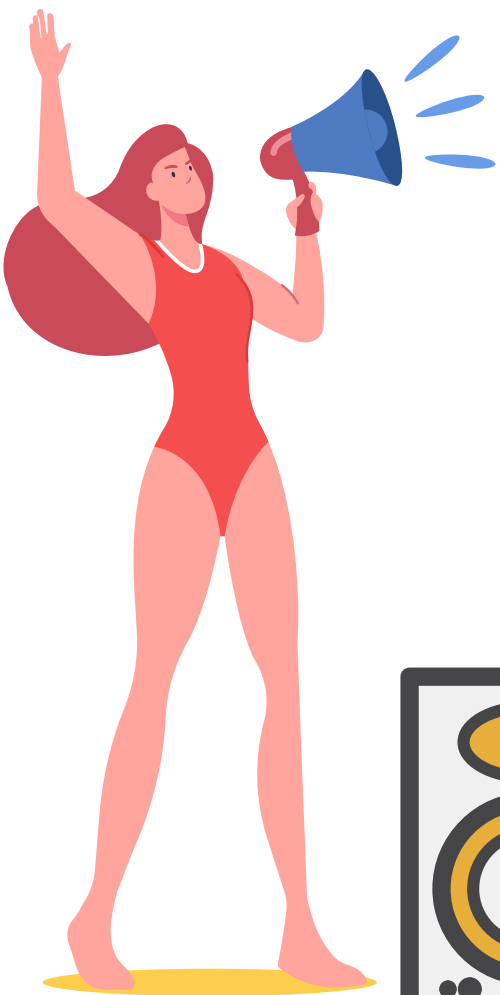
OVERVIEW 🔗

[bbc.co.uk/accessibility/forproducts/policy/](https://www.bbc.co.uk/accessibility/forproducts/policy/)



3. 配音的特点和翻译要求

配音翻译的特点



配音演员的选择

声音造型艺术（情绪、语气、音调、节奏）

录音场所和音像设备选择

与背景音乐的协调处理

配音翻译的基本要求

整体性：

- 配音应对照节目视频完成，配音的音色、音量、语速、情绪等与原节目相符合，可进行艺术性再创作。

本土化：

- 配音宜进行本土化配音。本土化配音应由以目标语言为母语的配音导演组织实施。根据需要，可配备同时掌握中文与目标语言的语言导演，协助母语配音导演准确解读台词、微调台词等。

配音与原声搭配：

- 纪录片、新闻等节目的旁白、解说应进行配音，人物同期声根据需要配音并保留部分原声。

对口型 (Lip-sync)

出发点:

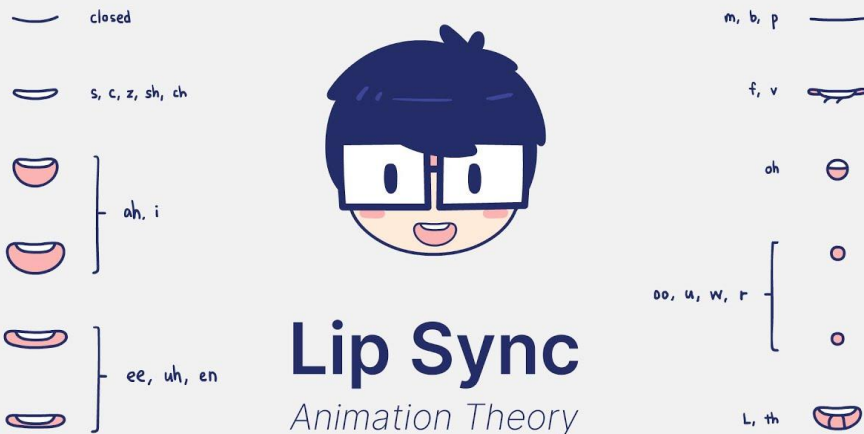
- 音素-视位对应理论 (Phoneme-Viseme Mapping)
- 人眼可容忍的声画延迟阈值为 $\pm 80\text{ms}$ (超过则产生“配音违和感”)

配音原则:

- 音节数匹配:
 - 原文 “I don't know” (3音节) \rightarrow 中文 “不知道” (3音节) 优于 “我不清楚” (4音节) 。
- 重音位置对齐:
 - 英语重音音节需对应中文逻辑重音 (如 “Science” \rightarrow “科学” 而非 “科学研究”) 。
- 爆破音可视性:
 - 保留 /p/、/b/ 等明显闭合嘴型的词汇 (如 “pop” 译 “啪” 而非 “爆裂”) 。

配音翻译实践:

- 3C原则: Content (内容)、Character (角色)、Culture (文化) 的优先级权衡 (Luyken et al., 1991) 。
- 观众对 Lip-sync 的敏感度排序: 儿童动画 > 真人剧情片 > 纪录片 (欧盟 HERMES 项目发现) 。



配音翻译的工作流程

传统精品流程

- 院线电影
- 3A游戏

AI参与的工业化流程

- 短视频
- 出海短剧
- 教育课件

1. 传统主流配音翻译流程（**Human-in-the-loop**）

素材准备

- 获取视频及分离的 **M&E Track**（音乐与音效轨）

剧本翻译与润色

- 听译/场记
- 口型适配

选角与录音

- 导演挑选配音演员（CV），进棚录音

混音与后期

- 将新的人声与 M&E 轨合成

2. AI 参与的成熟流程 (AI-Augmented)

AI 听写与打轴
(ASR & Spotting)

- **Subtitle Edit (Whisper 插件) 或 Arctime Pro**

AI 辅助翻译 (LLM
Translation)

- ChatGPT-4o/Gemini 3
- 现在的Prompt可以要求 AI: “请翻译这句话, 要求中文音节数与英文相近, 并且尽量匹配开闭口音。”

AI 预配音
(Reference Dub)

- Azure TTS / 剪映文本朗读

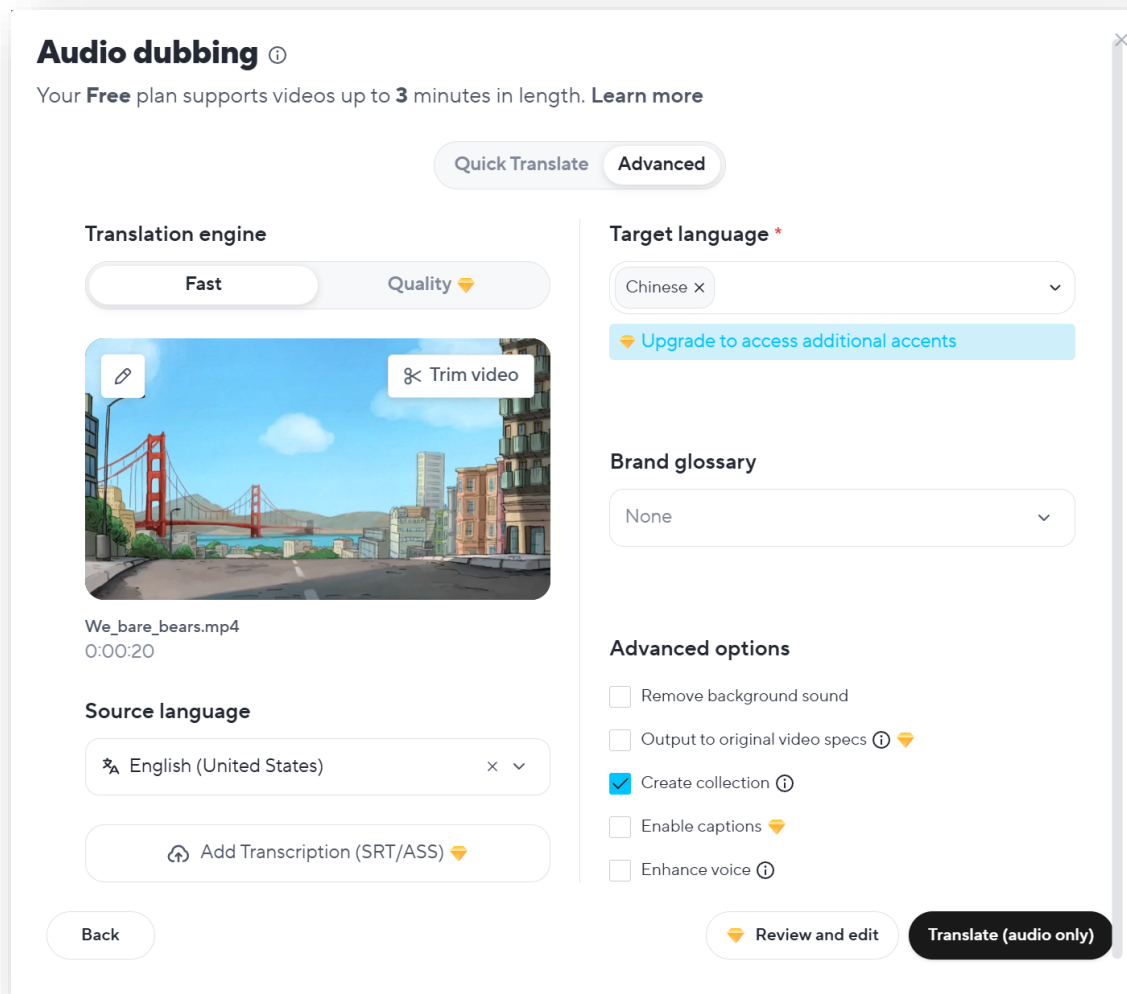
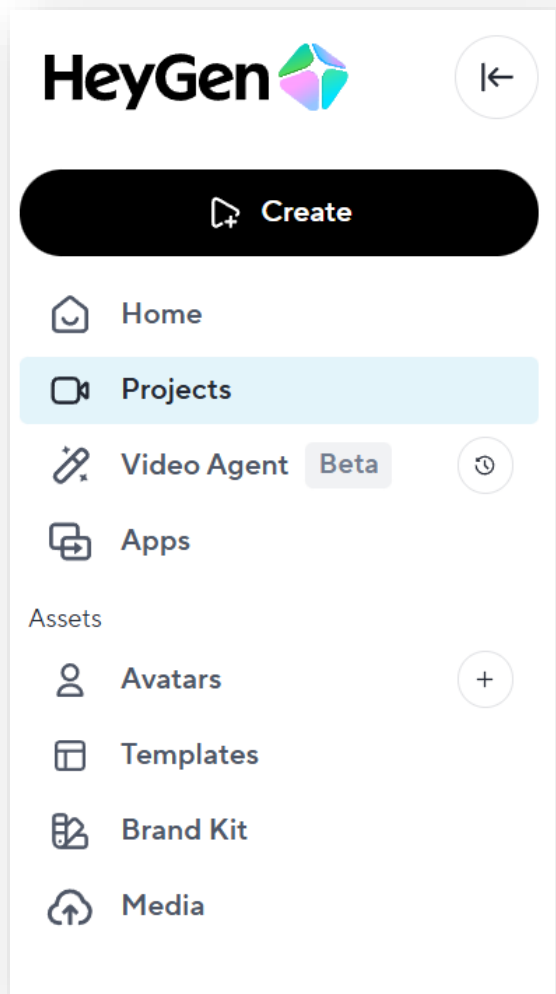
Have a try! 对口型工具 (lip-sync)

HeyGen (Video Translate)

D-ID

Rask.ai

工具：HeyGen 自动配音翻译



工具: Rask.ai

每年购买 Creator 和 Creator Pro 可享受 35% 折扣。价格即将变动 - 锁定所有计划的当前价格

02d:13h:52m:58s



PRODUCT HUNT
#1 Product of the Day

工具 ▾ 使用案例 ▾ 定价 资源 ▾ 博客

 Zh ▾ 登录

与销售人员交谈

免费试用

 受到 340 万用户的信赖  4,7 星级

用逼真的人工智能语音翻译视频

为全球企业提供由人工智能支持的音频和视频配音，可在应用程序内或通过应用程序接口使用

免费试用，无需信用卡

管理大量文件? [联系我们](#)

 EN  SP  JP  AR  ZH  FR  IT  HI



4.字幕翻译制作工具



字幕翻译制作软件（是否支持翻译记忆）

■ 支持翻译记忆的字幕软件

- SDL Trados Studio
- Trados Studio subtitling
- 人人译世界
- TransWAI
- VideoLocalize (云翻译)
- 其他

■ 不支持翻译记忆的字幕软件

- Aegisub
- Subtitle Workshop
- Subtitle Edit
- Arctime
- Time Machine
- 讯飞听见
- 网易见外
- 剪映

字幕翻译制作软件（本地 **or** 云端）

本地

- **Aegisub**
 - 开源、免费
 - 功能较为全面
- **Subtitle Edit**
 - 开源、免费
- **WinCaps Q4**
 - 付费、专业
 - 专业字幕编辑
- **Adobe Premiere Pro**
 - 专业级视频编辑软件

云端

- **TransWAI**
 - 支持字幕翻译、协作
 - 支持翻译记忆库
 - 提供机器翻译API接口
- **OOONA**
 - 云端+本地
 - 支持机器翻译、协作
- **剪映**
 - 视频编辑应用
 - 可智能语音识别字幕
 - 翻译功能付费



5. 经典字幕翻译流程案例 - Aegisub

Aegisub



Aegisub
应用

* We Bare Bears _ Opening Theme.txt - Aegisub 3.2.2

文件(F) 编辑(E) 字幕(S) 计时(T) 视频(V) 音频(A) 自动化(U) 查看(W) 帮助(H)

0:00:03 04 05 06 07 08 09 10 11 12 13 14 15 16 17

0:00:00.00 0:00:00.00 0:00:00.00 0 0 0

☐ 注释(C) Default 结束延迟后所选择(V) 特效 15

Yeah, we'll be there!

#	开始时间	结束时间	字/秒	样式	文本
1	0:00:00.00	0:00:00.00		Default	DA, DA, DA-BA-DA
2	0:00:00.00	0:00:00.00		Default	DA, DA, BA-DA-BA-DA-BA
3	0:00:00.00	0:00:00.00		Default	LET'S GO!
4	0:00:00.00	0:00:00.00		Default	We'll be there!
5	0:00:00.00	0:00:00.00		Default	A wink and a smile and a great, old time!
6	0:00:00.00	0:00:00.00		Default	Yeah, we'll be there!
7	0:00:00.00	0:00:00.00		Default	Wherever we are,
8	0:00:00.00	0:00:00.00		Default	There's some fun to be found!
9	0:00:00.00	0:00:00.00		Default	We'll be there when you turn that corner,
10	0:00:00.00	0:00:00.00		Default	We'll jump out the bush,
11	0:00:00.00	0:00:00.00		Default	With a big bear hug and a smile!
12	0:00:00.00	0:00:00.00		Default	We'll be there!

先打开字幕，
再打开视频

Aegisub视频字幕操作步骤



Aegisub视频字幕操作技巧

1. 关闭频谱分析模式，用空格键播放当前字幕
2. 确定该句大致区间，单击鼠标左键标定字幕出现时间（红色竖线）、单击鼠标右键标定字幕结束时间（蓝色线条）
3. 在英文输入法下，单击键盘上的G切换到下一行字幕

拓展思考 (Ask CAT 2.0 Tutor!)

如何制作双语字幕?

过程中使用什么工具? workflows如何?

如何制作双语软字幕? 硬字幕?

助教在此!



1. 使用 \N 换行符来增加中文字幕 (N要大写)
2. 提供单语字幕 (带时间轴), 让LLM制作
双语字幕

6. AI+字幕翻译流程



AI+案例1：视频字幕翻译

要求

- 你收到MP4格式的英文视频文件（会议）、需要添加中文字幕。
- 提交文件包括：中文字幕文件SRT，包含中文字幕的视频文件。

方法：

剪映

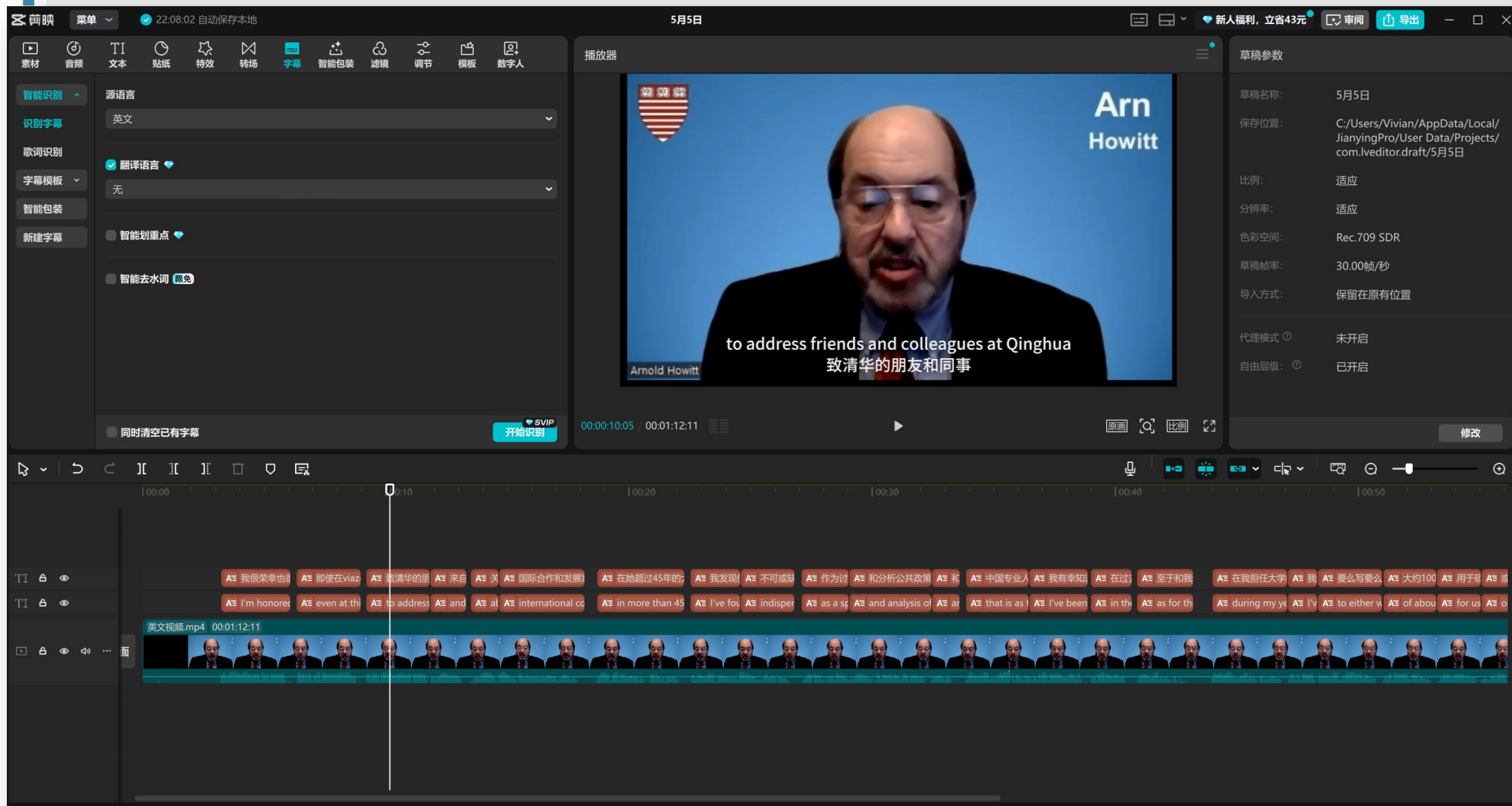
人人译视界

TransWAI

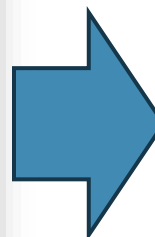
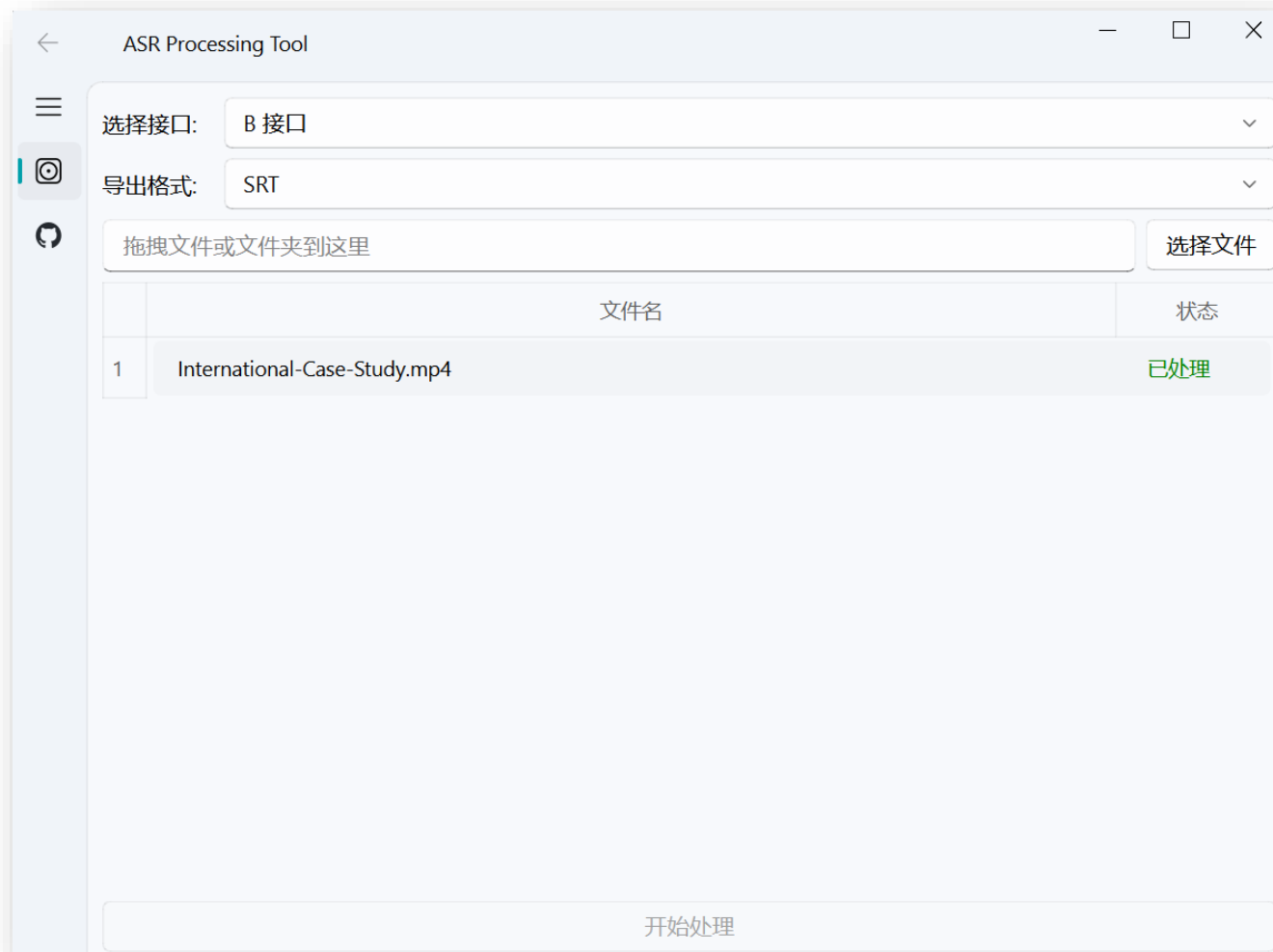
Asrtools




Microsoft
Clipchamp

方法1：剪映



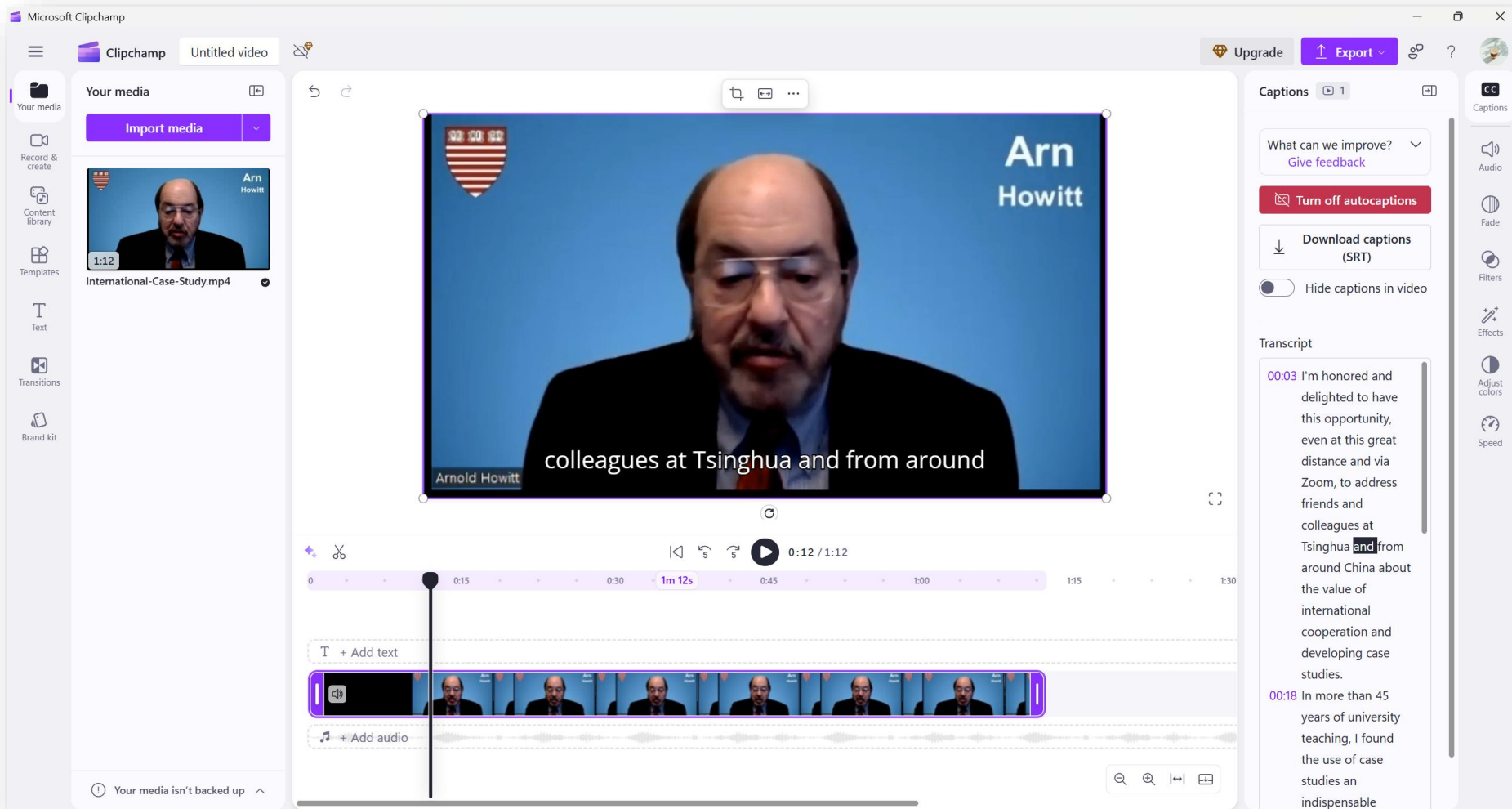
方法2: AsrTools



名称	修改日期
 International-Case-Study.mp3	5/5/2025 10:15 PM
 International-Case-Study.mp4	5/1/2022 2:30 PM
 International-Case-Study.srt	5/5/2025 10:15 PM

可快速抓取字幕和音频
(质量一般, 就是快!)

方法3: Microsoft Clipchamp



可识别字幕，
英文识别出色，
免费

方法4：人人译视界

The screenshot displays the 人人译视界 (Renyi) software interface, which is used for video editing and subtitle management. The interface is divided into several sections:

- Top Bar:** Contains tabs for '开始' (Start), '字幕编辑' (Subtitle Editing), '工作台' (Workspace), '文件队列' (File Queue), '视频压制' (Video Compression), '小工具' (Tools), and '设置' (Settings). It also includes a '系统消息' (System Message) button on the right.
- Toolbar:** A row of icons for various functions such as '新建' (New), '打开' (Open), '保存' (Save), '撤销' (Undo), '重做' (Redo), '关闭' (Close), '打开字幕' (Open Subtitles), '追加字幕' (Append Subtitles), '追加文本' (Append Text), '空白视频' (Blank Video), '打开视频' (Open Video), '关闭视频' (Close Video), '我的任务' (My Tasks), '创建项目' (Create Project), '存入云端' (Save to Cloud), and '云端文件' (Cloud Files).
- Main Video Player:** Displays a video of a man speaking. The video title is 'Arn Howitt'. The video player includes a progress bar, a volume control, and a '118/1809' indicator.
- Subtitle Editor:** A section for editing subtitles. It includes a '离线任务' (Offline Task) tab. The subtitle list shows a single subtitle at 00:00:04,720 with the text 'I'm honored and delighted to have this opportunity'. The editor includes a '查询' (Search) button, a '更改大小' (Change Size) button, and a '划词查询' (Highlight Search) button. It also has a '简繁转换' (Simplified/Traditional Conversion) button, an '识别画面字幕' (Identify Video Subtitles) button, and buttons for '智能听译' (Smart Listening/Transcription), '智能打轴' (Smart Timing), '智能配音' (Smart Dubbing), and '机器翻译' (Machine Translation).
- Subtitle List:** A table showing a list of subtitles. The first subtitle is highlighted in blue.
- Subtitle List Table:**

序号	开始时间	结束时间	字幕内容
1	00:00:04,720	00:00:08,140	I'm honored
2	00:00:08,140	00:00:10,980	Even at this
3	00:00:10,980	00:00:15,350	to address fri
4	00:00:15,350	00:00:20,310	about the val

- Bottom Bar:** Displays the original language (English) and the translated language (Simplified Chinese). It also shows the video duration (00:01:12,424) and the subtitle duration (00:01:08,850).

可压制嵌入式
字幕
(硬字幕)

<https://www.1sj.tv/html/support/download.html>

方法5: TransWAI 字幕翻译

International-Case-Study.mp4

自动保存: 2025-05-05 22:22:48

保存编辑

撤销

恢复

字幕检查

查找替换

快捷键

导入字幕

视频配音

下载字幕

导出视频

T 字幕

排版样式

旁白

导出列表

隐藏译文

自动翻译

自动播放

序号	时间码	识别结果
3	- 00:08.760 +	to address friends and colleagues at Tsinghua and from around China about the value of international cooperation and developing case studies and more than 45 years of university teaching.
	- 00:20.980 +	向清华和来自中国各地的朋友和同事介绍国际合作和发展案例研究的价值, 以及45年多的大学教学经验
4	- 00:21.880 +	I found the use of case studies and indispensable curricular tool as a spark for discussion and analysis of critical issues and public policy.
	- 00:32.240 +	我发现使用案例研究和不可或缺的课程工具作为讨论和分析关键问题和公共政策的火花
5	- 00:32.470 +	And governance.
	- 00:33.280 +	和治理

Arn Howitt

I found the use of case studies and indispensable curricular tool as a spark for discussion and analysis of critical issues and public policy. 我发现使用案例研究和不可或缺的课程工具作为讨论和分析关键问题和公共政策的火花

00:00:2500:01:121.0x

00:00:2200:00:2300:00:2400:00:2500:00:2600:00:2700:00:2800:00:2900:00:3000:00:3100:00:3200:00:3300:00:3400:00:3500:00:3600:00:3700:00:3800:00:3900:00:4000:00:4100:00:42

I found the use of case studies and indispensable curricular tool as a spark for discussion and analysis of critical issues and...

我发现使用案例研究和不可或缺的课程工具作为讨论和分析关键问题和公共政策的火花

And ... That is true for the Chinese professionals I've been privileged to know and teach ... As for the Americans I've worked with c

和治理 这对我在过去20年里有幸认识和教过的中国专业人士来说是真的 至于我在大学教职员工中与之共事的美国人

可配音
(2分钟11元)

AI+案例2：配音翻译

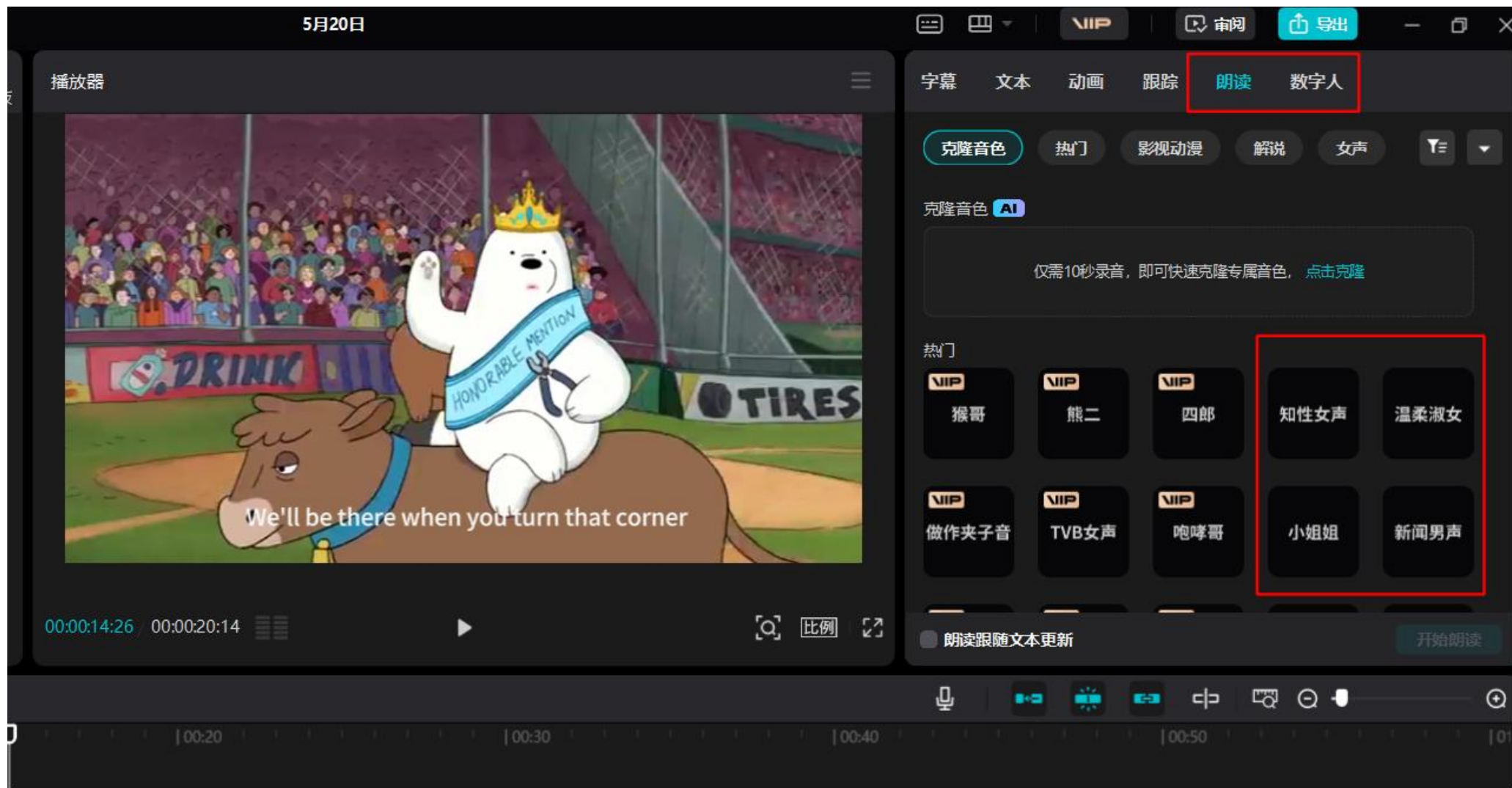
□ 要求

- 你收到 MP4 格式的英文视频文件、原轨的背景音乐+音效文件，需要添加中文配音。
- 提交文件包括：中文字幕文件 SRT，包含中文配音的视频文件

□ 思考

- 怎么样可以制作出中文歌曲配音？

剪映-字幕转配音



Rask.ai 视频译制配音

The screenshot displays the Rask.ai web interface for video dubbing. The top navigation bar includes a back arrow, the project name 'We_bare_bears', and buttons for 'View plans', 'Advanced tools', search, help, user profile, and 'Share'. The main interface is divided into two primary sections: a text/audio editing panel on the left and a video preview panel on the right.

Text/Audio Editing Panel:

- Language Selection:** A dropdown menu shows 'English' (with a US flag) and 'Chinese (CN) CN' (with a Chinese flag). A '+ New language' button is also present.
- Audio Segments:** Three segments are listed, each with a speaker selection (Speaker 1), a time range, and a 'Voice Clone' icon.
 - Segment 1: Time 00:00:00,008 — 00:00:01,587. Original text: 'Da-da-da-ra-da-da'. Translated text: '哒哒哒哒哒哒'.
 - Segment 2: Time 00:00:01,994 — 00:00:10,568. Original text: '-da-ra-da-da-da-da-da-da-da-da-da Ready, set, we'll be there To wink and smile with a big old smile We'll be there'. Translated text: '准备好了，我们马上就到，带着大大的微笑眨眨眼，我们会到的'.
 - Segment 3: Time 00:00:11,061 — 00:00:19,685. Original text: 'Wherever we are there's something fun to be found We'll be there when you turn that corner We'll jump out the porch with a big big organ smile We'll be there'. Translated text: '无论在哪，都有乐子，转角我们就在，带着大笑跳出来，我们会到的'.

Video Preview Panel:

- Video Player:** Shows the original video with three characters: a boy, a panda, and a bear. The video is labeled 'TRANSLATED BY RASK'. The player includes a progress bar (00:00:04,199 / 00:00:20,015), a volume icon, and a 'CC' icon.
- Controls:** A large purple 'Apply changes' button is prominent. Below it are buttons for 'Lip-sync New!' and 'Download'.
- Footer:** A small text note states 'Voice Synthesis powered by ElevenLabs'.

Timeline: A horizontal timeline at the bottom shows the audio segments aligned with the video. It includes a volume slider and a 'Speaker 1' label for each segment.

配音原理：自动语音识别 -> AI 翻译 -> 声纹克隆 -> 视觉口型矫正 -> 自动混音

拓展问题

如何实现智能的翻译 → 配音 → 歌曲配音流程？

 **Synthesizer V Studio 2 PRO**



资源推荐

视听AI 工具推荐

环节	免费/开源/学生友好	专业/付费/商业级
听写 (ASR)	Subtitle Edit (内置Whisper)	Otter.ai
翻译 (MT)	DeepL (免费版), ChatGPT-4o	MemoQ + OpenAI插件
配音 (TTS)	Edge TTS, GPT-SoVITS	ElevenLabs
视频生成	SadTalker (HuggingFace)	HeyGen, Rask.ai

END