

國立臺灣大學文學院翻譯碩士學位學程

碩士論文

Graduate Program in Translation and Interpretation

College of Liberal Arts

National Taiwan University

Master's Thesis

視聽翻譯與媒體近用：

電影《她們》隱藏式字幕的多模態研究

Audiovisual Translation and Media Accessibility:

A Multimodal Study of Closed Captions of *Little Women*

黃瑜安

Yu-An Huang

指導教授：孔思文博士

Advisor: Szu-Wen Kung Ph.D.

中華民國 113 年 6 月

June 2024



國立臺灣大學碩士學位論文
口試委員會審定書
MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY



視聽翻譯與媒體近用：電影《她們》隱藏式字幕的多模態研究

Audiovisual Translation and Media Accessibility:

A Multimodal Study of Closed Captions of *Little Women*

本論文係黃瑜安（R09147005）在國立臺灣大學翻譯碩士學位學程完成之碩士學位論文，於民國 113 年 6 月 24 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Graduate Program in Translation and Interpretation on June 24, 2024, have examined a Master's Thesis entitled above presented by YU-AN HUANG (R09147005) and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

孔思文

（指導教授 Advisor）

陳嘉儀

李根芳

Acknowledgements



I would like to express my sincere gratitude to my thesis advisor, Professor Szu-Wen Kung, for her invaluable insights and tremendous support throughout the writing process. This study would not have been possible without Professor Kung's encouragement and open-minded guidance.

As a graduate student, I am fortunate to learn from Professor Ken-Fang Lee in the Cinema and Translation class, where I expanded my knowledge of film studies and translation theories. Professor Chia-Chien Chen's Localization in Audiovisual Translation class helped me unravel the intricacies of subtitle translation and the tricks of the trade. These courses played a vital role in advancing my research.

Many thanks to the Senior Specialist, Ms. Yu-Ting Li, who constantly offered timely assistance with any administrative matters.

I am honored to have met many kind and talented friends in the graduate program, and I will always remember our time at the BRAFS thesis support group, where we shared our (blood), sweat, and tears while writing our theses.

Lastly, I am thankful for my dear family, who unconditionally supported me through my learning journey and afforded me the privilege to learn and research.

Abstract



With the rapid growth of online streaming services, closed captions serve as an essential form of audiovisual translation for the deaf and hard-of-hearing, conveying comprehensive auditory elements like sound effects, music, voice quality, volume, and tone. This research reviews the development of audiovisual translation, the crucial concept of media accessibility for diverse audiences, and the connection between multimodality and translation. Aiming to investigate the complex multimodal meaning-making in audiovisual content, analyze the intersemiotic translation strategies employed, and explore captioning practices that promote media accessibility, this study employs a multimodal framework to analyze the captioning of the film *Little Women* (2019). The analysis reveals the interplay of modes such as language, image, sound, and music, where captioners utilize strategies such as adding speaker identification, adjusting caption positioning, and using icons and typographical formatting to provide plot-relevant details for the deaf and hard-of-hearing. Through this study, the researcher hopes to offer diverse perspectives on subtitling practices in Taiwan, ensuring that individuals with varying needs can immerse themselves and enjoy audiovisual works.

Keywords: Audiovisual Translation, Closed Caption, Deaf and Hard-of-Hearing, Media Accessibility, Multimodality

摘要



隨著線上串流平臺興起，隱藏式字幕（Closed Captions）成為協助聽覺障礙者的重要視聽翻譯形式，能夠傳達音效、音樂、音質、音量與語調等完整的聽覺元素。本研究回顧了視聽翻譯的發展、面向多元觀眾的媒體近用概念以及多模態與翻譯之間的連結。本研究採用多模態的理論框架，分析 2019 年電影《她們》（*Little Women*）的隱藏式字幕，旨在探究視聽作品中複雜的多模態意義建構、分析其中運用的符際翻譯策略，並探討能促進媒體近用的字幕翻譯實務。研究結果揭示了語言、影像、音效與音樂等模態之間的交互作用，以及字幕譯者透過點出說話者身份、調整字幕位置、活用符號和排版格式等策略，為聽覺障礙者提供與劇情相關的聲音細節。希望透過這項研究，為台灣的字幕服務提供多元的視角，確保具備不同需求的族群都能沉浸在視聽作品之中，盡情享受其樂趣。

關鍵詞：視聽翻譯、隱藏式字幕、聽覺障礙者、媒體近用、多模態

Contents



Master's Thesis Acceptance Certificate	i
Acknowledgements	ii
Abstract	iii
摘要	iv
1. Introduction	1
2. Literature Review	8
2.1 Audiovisual Translation	8
2.2 Media Accessibility	18
2.3 Multimodality and Translation	24
3. Research Methods	32
3.1 Film Text Selection: <i>Little Women</i> (2019)	32
3.2 Justification of Text Selection	35
3.3 An Overview of Netflix's Style Guide	38
4. Results and Analysis	45
4.1 Categorization	46
4.2 Attribution	54
4.3 Explanation	63
4.4 Symbols and Icons	69
4.5 Typography	71
4.6 Discussion and Implications	76
5. Conclusion	83
References	87
Appendix: Netflix's Style Guide for Subtitles for the Deaf and Hard of Hearing	95

List of Tables



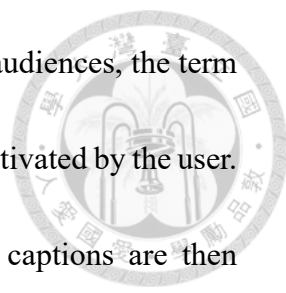
Table 1 Transcription of closed captions	44
Table 2 Transcription of Case 1: Categorization, exhibiting the sound resulting from the action	46
Table 3 Transcription of Case 2: Categorization, identifying speakers	48
Table 4 Transcription of Case 3: Categorization, depicting the action that produces the sound	50
Table 5 Transcription of Case 4: Categorization, presenting background dialogue	52
Table 6 Transcription of Case 5: Categorization, describing animal sounds.....	53
Table 7 Transcription of Case 6: Attribution, assigning qualities to a sound	55
Table 8 Transcription of Case 7: Attribution, describing the tone of voice.....	56
Table 9 Transcription of Case 8: Attribution, describing the tone of voice.....	58
Table 10 Transcription of Case 9: Attribution, identifying the languages.....	59
Table 11 Transcription of Case 10: Attribution, giving details to music in films.	62
Table 12 Transcription of Case 11: Explanation, delineating the production of a sound and its relations to other entities	64
Table 13 Transcription of Case 12: Explanation, delineating the production of a sound and its relations to other entities	67
Table 14 Transcription of Case 13: Symbols and Icons	69
Table 15 Transcription of Case 14: Typography, adjusting caption positioning ..	71
Table 16 Transcription of Case 15: Typography, adjusting caption positioning ..	73
Table 17 Transcription of Case 16: Typography, italicizing captions.....	74

1. Introduction



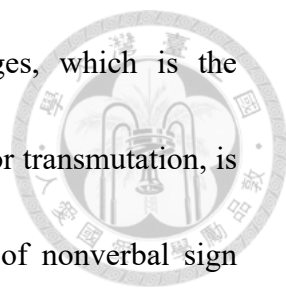
As the COVID-19 pandemic led to social distancing, home quarantine, and isolation, it reshaped the media and entertainment landscape. Nowadays, fast-growing streaming services like Netflix, Disney+, Hulu, and Apple TV have become a more convenient and affordable way to consume a variety of media content, such as films, TV series, and reality shows. When people have access to these works from all around the world, it is clear that translation plays a vital part in breaking cross-cultural barriers and formulating meanings for the target audiences. In this case, subtitles are one of the most common tools for global audiences to better understand international audiovisual work. Subtitles, originated from intertitles in silent films, are written texts superimposed on audiovisual contents that deliver the meaning of the source speech in different target language versions, which appears in synchronization with the spoken content in the film (Perez-Gonzalez, 2014).

When it comes to watching films, we tend to assume that all audiences are hearing viewers, and subtitles are usually designed for these audiences in their target languages. But what if the audiences have special needs? While traditional subtitles are created for hearing audiences, “closed captions (CC)” or “subtitles for the deaf and (the) hard-of-hearing (SDH or SDHH)” provide media access for hearing-impaired viewers.



Compared to subtitles embedded in the video that are visible to all audiences, the term “closed captions” indicates that captions are usually “closed” until activated by the user. By turning on the remote control or selecting the menu option, captions are then displayed on the screen (Szarkowska, 2020: 249). In terms of terminology, Neves (2008: 130) argues that “closed captions” are generally used in the U.S., Canada, and Australia, while “SDH” is more common in European countries, and the terminological differences between closed captions and SDH can be referred to as the differences in British and American English, such as “flat/apartment” or “taxi/cab” varieties. Similarly, Vanderplank (2016: 9) notes that different terms are used to describe the same-language subtitles designed for the deaf and hard-of-hearing people: “closed captions” in North America; “teletext subtitles” in the UK and some other countries; and lastly, “subtitles for the deaf and hard-of-hearing” in non-English-speaking countries. Whether closed captions or SDH, they both convey all auditory elements including music, sound effects, intonation, speaker identification, and other non-speech factors (Zárate, 2021: 4), which are mostly written in the source language of the media content.

In his influential taxonomy of translation, Jakobson (2000: 114) distinguished three types of translation as intralingual translation, interlingual translation, and intersemiotic translation. Intralingual translation, also known as rewording, means translation within the same language, which can be understood as paraphrasing (Whyatt, 2017: 176).

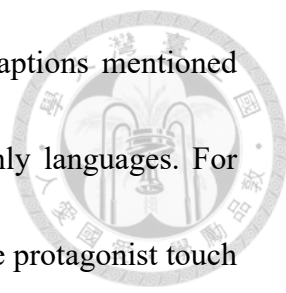


Interlingual translation is translation between different languages, which is the translation proper (Jakobson, 2000); and intersemiotic translation, or transmutation, is defined as “an interpretation of verbal signs by means of signs of nonverbal sign systems” (Jakobson, 2000: 114), such as films adapted from novels. In light of the classification, subtitling can be categorized into interlingual subtitling and intralingual subtitling (Diaz Cintas & Anderman, 2009; Vanderplank 2016; Szarkowska 2020; Zárte 2021). Interlingual subtitling involves the classic translation activity from the source language to the target language, for example, the Chinese translation of an English film. On the other hand, intralingual translation refers to the translation that happens within the same language (Diaz Cintas & Anderman, 2009:12), such as the English SDH for an American TV series. Closed captions and SDH are usually viewed as intralingual subtitling since cases of closed captions or SDH of foreign films are still uncommon these days, for people generally think that interlingual subtitling can fulfill the needs of the deaf and hard-of-hearing (Szarkowska, 2013). The discussion here reflects the fact that most interlingual subtitles are not tailored for the hearing impaired (Zárte, 2021: 5), which only contain limited information that leaves out the non-speech elements. The information provided by the subtitles failed to empower the deaf and hard-of-hearing audience to grasp the whole meaning of the film and immerse themselves in the storyline. In this research, we want to pay attention to the diverse

needs of audiences, especially the hearing-impaired viewers, and to investigate the styles and strategies of closed captions.

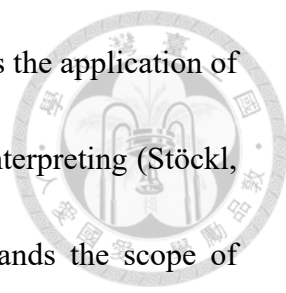


With different audiences in mind, interlingual subtitles and closed captions are presented distinctively in the meaning-making process. Due to the spatial and temporal considerations, interlingual subtitles offer more condensed and streamlined versions of the original dialogue in the film text (Perez-Gonzalez, 2014). For instance, redundancy of hesitation and interjections, such as *ouch*, *wow*, *ugh*, are usually omitted in the subtitles. By observing the intonation, facial expressions, and gestures of the characters, hearing viewers can infer the meaning without the help of subtitles. On the other hand, closed captions provide more detailed sound features in the film, which are important for the development of plots or the creation of the atmosphere, for the hearing-impaired viewers (Diaz Cintas, 2020). From closed captions, the hearing-impaired can “see” the sounds in written words and imagine the atmosphere created by a variety of sounds in the film, such as chattering in the shop, a baby crying, or a man knocking on the door, etc. In addition, closed captions sometimes include the name of the speakers and can be presented in changeable positions to the left or right of the screen if it is important to identify the speakers or to indicate the direction of an off-screen sound (Diaz Cintas, 2020: 160). Regarding the differences between interlingual subtitles and closed captions, more details will be explored in 2.1.



From the characteristics of interlingual subtitles and closed captions mentioned earlier, we can see that meaning-making in films involves not only languages. For example, in the film, *Little Women* (2019), in a scene we can see the protagonist touch her hands nervously, yell at her sister with anger, and whisper her thoughts to herself. These gestures and different tones of voice all play a role in establishing meanings in the film. Both the visual and auditory elements, like texts, images, music, and sound effects, influence the meaning-making process. Therefore, multimodality was introduced as an interdisciplinary approach that investigates the combination of multiple modes that constitute meanings. Kress and Van Leeuwen (2001: 20) defined “multimodality” as “the use of several semiotic modes in the design of a semiotic product or event.” And the crucial term “mode” is defined as “semiotic resources which allow the simultaneous realisation of discourses and types of (inter)action” (Kress & Van Leeuwen, 2001: 21). The multimodal framework provides a holistic view to exploring the verbal and non-verbal elements during the translation process and shifts the focus from the centrality of “language” to the emphasis on “meaning” (Kress, 2020: 24).

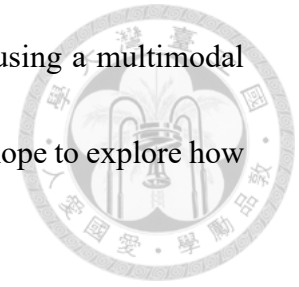
Among the existing research, there are two main ways that translation studies have interacted with multimodality (Zhang & Feng, 2021: 8). One is the development of theoretical and methodological frameworks based on insights from multimodal



discourse analysis (Munday, 2012; Kaindl, 2013); the other involves the application of multimodal approaches to study the translated texts and dialogue interpreting (Stöckl, 2004; Pérez-González, 2014). The concept of multimodality expands the scope of translation studies and enables us to see translation from a linguistic frame to a semiotic frame (Kress, 2020: 24). In other words, multimodal analysis allows us to look into different modalities in the communication process, and when the modalities are restricted for the hearing-impaired viewers, closed captions can provide them with a way to compensate for the lack of sound modality. Through the concept of multimodality, the research of closed captions/SDH no longer focuses solely on linguistic or technical aspects. Incorporating multimodal approaches in closed captions/SDH studies provides a more comprehensive understanding of how meaning is constructed and presented through various modes, which can help improve the accessibility of closed captions/SDH.

In this research, we will look into the interactions of subtitles and the film text and observe how subtitles create meanings together with auditory and visual elements in the film text. Even though typically classified as intralingual translation, closed captions/SDH also involve intersemiotic translation strategies that convert speech and sound effects into languages. Since we focus on deaf and hard-of-hearing viewers who cannot (wholly) perceive one of the key semiotic modalities, sound mode, we will

analyze the intersemiotic translation strategies of closed captions using a multimodal framework to uncover the process of compensation. In the end, we hope to explore how to improve media access for viewers with hearing loss.



The thesis comprises five sections. The first introduces the research background and points out foundational concepts regarding closed captions/SDH and multimodality. The next section provides an overview of existing literature on audiovisual translation, with particular attention to subtitle translation, discussions around closed captions and media accessibility, as well as key research on translation and multimodality. The third component outlines the research questions, justifies the selection of the film text, introduces the general style guide for captioning, and elaborates on the data collection and classification approaches. The fourth section presents and analyzes cases of closed captions, categorized and discussed from a multimodal framework. Lastly, the conclusion section summarizes the primary findings of this research, acknowledges its limitations, and offers recommendations for future endeavors.

2. Literature Review

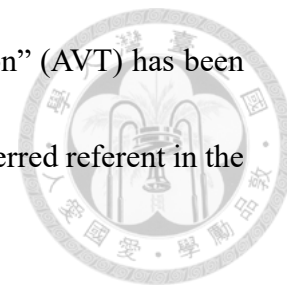


To understand the origin and development of subtitles, this section will first give an overview of audiovisual translation and three categories of audiovisual translation: subtitling, revoicing, and assistive forms of audiovisual translation. As we look into the role of audiovisual translation, it is clear that “media accessibility” has become a pivotal topic for translation nowadays. In the paragraph that follows, the concept of media accessibility allows us to reflect on whether we excluded people with different needs from enjoying audiovisual products and how we can strive to create accessibility for all. Lastly, we will shed some light on multimodality and Stöckl’s categorization of four core modes to better understand the intersemiotic translation strategies.

2.1 Audiovisual Translation

Originally conceived as a way to distribute films across countries, “audiovisual translation” (AVT) is now used to mediate a wide range of screen-mediated texts (Pérez-González, 2014: 12). As audiovisual translation started to touch upon wide-ranging fields and issues, many terms have been used to identify the field of practice and research field, including “film translation”, “screen translation”, “language transfer”, “versioning”, “constrained translation”, and “(multi)media translation” (Pérez-González, 2014; Perego & Pacinotti, 2020). Amid the bloom of the terminology,

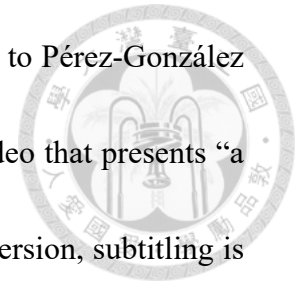
Díaz Cintas and Remael noted that the term “audiovisual translation” (AVT) has been increasingly recognized and is rapidly establishing itself as the preferred referent in the field (2007: 12).



According to Gambier (2012: 47), audiovisual products or performances “consist of a number of signifying codes that operate simultaneously in the production of meaning.” Both verbal and non-verbal signs are used in the meaning-making process, and identifying the relationships between them has become a key issue in audiovisual translation studies. As a practice, a product, and a discipline, the history of audiovisual translation is connected to the origin of films (Perego & Pacinotti, 2020). In 1895, the French Lumière brothers, deemed as the very first filmmakers, held the first commercial screening in history, and the groundbreaking invention raised the curtain on the future of cinema. From silent films to the first full-length feature film with synchronized dialogue, *The Jazz Singer* (1927), cinema was perceived as an art form and a business, and questions of how to export films and overcome language barriers were quickly raised (Gambier, 2012). To make audiovisual products comprehensible to audiences speaking different languages, dubbing and captioning/subtitling are common approaches adopted by different countries.

Pérez-González (2014: 12) pointed out three major types of audiovisual translation: subtitling, revoicing (like dubbing and voice-over), and assistive forms of audiovisual

translation (such as SDH and audio-description). Firstly, according to Pérez-González (2014: 15-16), subtitles are fragments of the written text on the video that presents “a target language version of the source speech.” In a more detailed version, subtitling is defined by Diaz Cintas and Remael (2007: 8) as:



A translation practice that consists of presenting a written text, generally on the lower part of the screen, that endeavors to recount the original dialogue of the speakers, as well as the discursive elements that appear in the image (letters, inserts, graffiti, inscriptions, placards, and the like), and the information that is contained on the soundtrack (songs, voices off).

Compared to Pérez-González's definition, Diaz Cintas and Remael point out the translation of non-speech elements in the film. Elements in the image and information on the soundtrack are included in subtitling, reflecting the multimodal nature of subtitles. Subtitling involves the shift from the oral to the written mode in which the subtitles (target text) are presented together with the audiovisual text (source text). Due to the spatial and temporal constraints, specific guidelines are created to establish a set of standards for professional subtitling.

In some countries, subtitles are mainly created to serve people with loss of hearing,

which combines features of subtitling and assistive forms of audiovisual translation.

For instance, according to BBC's Subtitle Guidelines (2022), the primary aim of subtitles (also known as closed captions here) is to assist viewers with loss of hearing,

but they are actually used by a wide range of people, and most of the viewers using

subtitles are not hard-of-hearing. The guidelines state that subtitles need to be faithful

to the audio content; nevertheless, translators must balance this against other factors

such as the character's action on the screen, the pace of speech, and the visual elements.

For example, when a character is speaking rapidly, the translator must consider the

following questions: Are viewers having trouble reading the subtitles at the rate of

speech? Is it necessary to edit out some words to meet the reading speed? If there are

shot changes during the conversation, should the subtitles be coordinated with those?

How to place the subtitles to avoid blocking the important visual elements on screen?

In general, subtitles prefer verbatim speech that gives the hearing-impaired more access

to the soundtrack. Words like "so", "but", "well", or "you know" are preserved in

subtitles for deaf or hard-of-hearing viewers and lip-readers. When it is necessary to

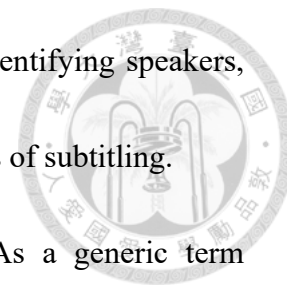
streamline the subtitles, it is better to evenly delete a bit of every sentence rather than

eliminate the entire sentence. In the BBC guidelines, each line is limited to 37

characters with a maximum of two lines as recommended. If necessary, three lines can

be used if no essential visual elements are covered. In addition to editing the text, rules

of matching subtitles to speaking speed and shots, positioning, identifying speakers, describing sound effects, and so on are all covered in the guidelines of subtitling.



The second type of audiovisual translation is revoicing. As a generic term consisting of several spoken translation methods, “revoicing” comes in various forms, such as dubbing, voice-over, simultaneous interpreting, free commentary, and narration (Pérez-González, 2014). Generally, the original soundtrack is replaced by the soundtrack in the target language, and the revoiced soundtrack must be synchronized with the video content. However, in voiced-over programs, the original and revoiced soundtracks exist at the same time. The slightly audible original soundtrack can create a sense of authenticity, which is why voice-over is used in documentaries and interviews. Revoicing allows viewers to enjoy the video in their native language without reading subtitles, but it can be a challenge for dubbers because they have to match the lip movements of characters, pay attention to the tone of their voices, etc.

Lastly, with growing concerns about social inclusion, people start to pay attention to the different needs of particular demographics, such as people with sensory disabilities. As a consequence, assistive forms of audiovisual translation like SDH (or closed captions), respeaking, and audio description aim to provide opportunities to access information provided by the media (Pérez-González, 2014: 24). For instance, SDH offers a text with both verbal and non-verbal renditions (Gambier, 2012: 49) that

includes all the sound features in the material, which makes it different from traditional subtitling. SDH utilizes punctuation marks, symbols, capital letters, labeling, and positioning to deliver the intonation, emotion, accents, and music in the video. By doing so, the deaf and hard-of-hearing can make sense of the nuances in the character's voice and the background sounds, which can help them better understand the plot and enhance the viewing experience. When the character stresses a certain word, it can be shown in caps as "It was HIS fault, not mine." When the character speaks in a particular accent, the subtitles can point it out as "[ENGLISH ACCENT] A cup of tea, please." These detailed features in SDH not only benefit the hearing-impaired but also those watching the video in noisy spaces or non-native speakers.

However, the concept, form, and technology of SDH/closed captions were not predetermined from the beginning but underwent several changes over time. As mentioned earlier, *The Jazz Singer* (1927) introduced the public to the fascinating world of "talkies", revolutionized the filmmaking business in Hollywood, and forever changed how we perceive films. From then on, motion pictures came with synchronized sound and later became standard in "talking pictures". Unfortunately, under this innovative transformation, the deaf and hard-of-hearing community were denied the right to enjoy the comprehensive cinematic experience. In 1947, as a deaf person himself, Emerson Romero created an approach that put "text-only frames" between

scenes, which is similar to intertitles in silent films (Neves, 2005). It provided a solution for the deaf and hard-of-hearing people, but the film became too long and the approach was not favored by Hollywood.



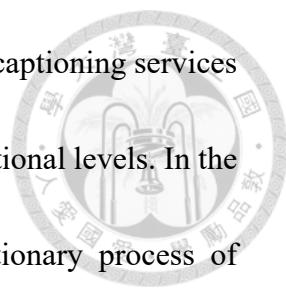
Later in 1949, film producer J. Arthur Rank devised a method that projected captions on another smaller screen in the bottom left corner of the main film screen; nevertheless, it was impossible for the audience to look at the two screens at the same time, but the idea was inspiring (Szarkowska, 2020; Stark, n.d.). Shortly after Rank's experiment, a new technique of captioning was created in Belgium, which involved etching captions directly onto a finished print of the film (Stark, n.d.). This groundbreaking invention led to the 1949 establishment of Captioned Films for the Deaf (CFD), a fund-raising nonprofit dedicated to making films accessible to the deaf audience (Szarkowska, 2020; Stark, n.d.). The organization utilized the Belgian technique and continued to persuade people in the film industry to caption their movies, and in 1951, the first open-captioned film, *America the Beautiful*, was born (Stark, n.d.). In 1958, with the passing of the Captioned Films Act (Public Law 85-905), CFD started to receive federal funding for open captioning, and the later Public Law 87-715 allowed CFD to acquire and caption educational films (Szarkowska, 2020; Stark, n.d.).

Szarkowska (2020) pointed out three significant milestones in 1972 that brought captioning into a new age: the first captioned TV program, *The French Chef*, aired on



Public Broadcasting Service (PBS); the Caption Center (now Media Access Group) at WGBH-TV, the first captioning agency, was founded in Boston; and BBC announced its new Ceefax service that provides teletext on screen for the deaf. In addition, as technology advances, “closed captions” are created as a flexible option that can be switched off when not needed. In the 1970s, the National Bureau of Standards (NBS) in the United States worked together with television services to develop encoding equipment and improvements to the captioning format, and in 1980, they started to transmit closed captions on selected programs (National Institute of Standards and Technology, 2016). Following the blooming development of captioning services, the National Captioning Institute was also established in 1980, aiming to caption TV programs and produce decoders that make captions appear on TV (Szarkowska, 2020; Stark, n.d.). From then on, captioning services have become more available around the world (Robson, 2004, as cited in Szarkowska, 2020: 257).

In Europe, the British followed in the footsteps of Americans and introduced the teletext system, similar to closed captioning in the US, on television in the 1970s (Neves, 2005). Later, the system was adopted by several countries across Europe, Asia, Africa, and the Pacific (Neves, 2005). Nowadays, TV programs, DVDs, Blu-ray discs, and streaming platforms offer assistive forms of audiovisual translation to provide access to hearing-impaired viewers.

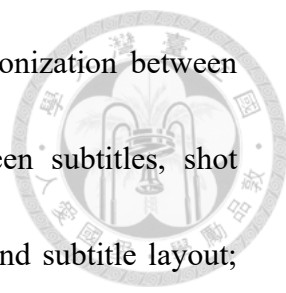


In addition to technological advancement, the development of captioning services is closely linked to legislative progress at both domestic and international levels. In the United States, a series of legislation helped promote the evolutionary process of captioning: the Captioned Film Act of 1958 authorized a loan service of captioned films for the deaf and raised public awareness of captioning through government regulations; the Americans with Disabilities Act of 1990 required public service announcements shall include closed captioning of the verbal content; the Television Decoder Circuitry Act of 1990 demanded television receivers to incorporate built-in decoder circuitry capable of displaying closed captioned television transmissions, and required the Federal Communications Commission (FCC) to ensure the continued availability of closed captioning services to viewers as new video technologies emerge (Described and Captioned Media Program, n.d.; U.S. Department of Justice Civil Rights Division, n.d.; National Association of the Deaf, n.d.).

At the international level, the UN Convention on the Rights of Persons with Disabilities was adopted in December 2006. The convention followed decades of effort by the United Nations to emphasize that individuals with disabilities of any kind are entitled to all human rights and fundamental freedoms as everyone else. It required participating states to adopt suitable measures to encourage the design, advancement, production, and dissemination of information and communications technologies and

systems that are accessible to all. As the front-runner of closed captioning services, the United States appointed the FCC to regulate the captioning services and set up captioning rules for deaf and hard-of-hearing viewers. As stated by the FCC, there are four captioning rules for all television programming: accuracy, synchronicity, completeness, and position. Captions should match the spoken words in dialogues and offer as many sound effects as possible; captions should appear simultaneously with the corresponding sound and display at a readable speed; captions should be presented from start to finish and be as complete as possible; captions should be properly placed without blocking the important visual messages (UN, 2006; FCC, 2021).

As subtitles/captions became popular practices in audiovisual translation, scholars started to investigate the factors that influence the translator's decision-making in subtitling and explored the relationship between subtitles and all the elements in the audiovisual text. Chen (2019: 5) discussed three crucial considerations of subtitling practice: the semiotic nature of subtitle translation, the overcoming of technical constraints, and the narrative flow to bridge the cultural gap. These considerations are indispensable in developing a theoretical framework for subtitle translation. Focusing on captioning and subtitling for deaf and hard-of-hearing audiences, Zárate (2021) indicated three considerations for captioning and subtitling: **linguistic considerations**, including text editing, segmentation and line-breaks, non-standard language, and



typographical details; **technical considerations**, including synchronization between subtitle and sound/image, reading speeds, minimum gap between subtitles, shot changes, subtitle file formats, the position of subtitle on screen, and subtitle layout; **special requirements**, including speaker identification, paralinguistic features, music, sound effects, and silence. Zárate's considerations will be explored from a multimodal perspective when analyzing the data of this study.

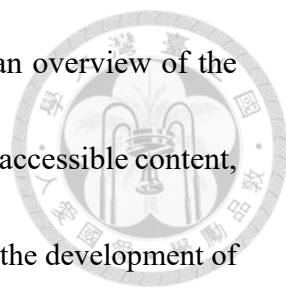
2.2 Media Accessibility

The core of audiovisual translation is to make audiovisual content accessible to a wider range of audiences. In an era where everything happens on screens, audiovisual texts have become an indispensable part of our daily life, which provides information, entertainment, and experiences that shape our minds. For the sensory-impaired community, access to media content gives them opportunities to enjoy cultural activities, performances, or services, such as cinemas, theatres, and museums. In addition to people born with sensory disabilities, as more countries become aging societies, the specific needs of elderly people are being discussed in an effort to create a world of true social inclusion (Perego & Pacinotti, 2020). Thus, the concept of “accessibility” came under the spotlight.

In fact, accessibility is not a completely new issue in audiovisual translation as in

2003 Gambier stated that “the key word in screen translation is now accessibility” (Gambier, 2003: 181). Nowadays, accessibility has become a vital concept in audiovisual translation, which has led to research on media accessibility that includes many socially relevant topics in our society (Greco & Jankowska, 2020). The study of media accessibility within audiovisual translation can be divided into four major categories: the history of audiovisual media accessibility; the development of accessibility legislation and provision; the textual and multimodal analysis of content-based audiovisual media accessibility services; and the recent experimental studies (Jankowska, 2020: 240).

To begin with, the history of audiovisual media accessibility can be traced back to the early days of cinema when intertitles were used to convey dialogue and narration in the image (Diaz Cintas & Remael, 2007). Since then, as the first sound film was screened, intertitles were replaced by a variety of solutions such as subtitling and dubbing. As the advent of new technologies reshaped the production, distribution, and consumption of audiovisual products, it also discriminated against and prevented certain people from accessing media services and products, which raised the issues of accessibility (Greco & Jankowska, 2020: 58). Over time, media accessibility started from a subdomain of audiovisual translation studies and emerged as an interdisciplinary field, which is accessibility studies (Greco 2018). Another topic of media accessibility studies is the



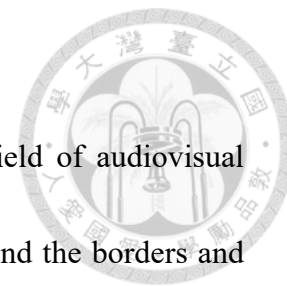
development of accessibility legislation and provision. It gives us an overview of the current legislation in different countries that ensures the provision of accessible content, such as the Convention on the Rights of Persons with Disabilities and the development of governmental regulations in America and Europe. As the third major category, the textual and multimodal analysis of content-based audiovisual media accessibility services explores a variety of accessible solutions, such as audio description, closed captions, or sign language interpretation, for specific audiovisual content, including but not limited to films, speeches, and even concerts. The final research type is the experimental studies, such as the research of users' reactions to different linguistic and technical criteria (Jankowska, 2020). By examining their reactions, we can gain a deeper understanding of diverse viewers' preferences and how to improve the quality of accessible services, creating a more friendly and inclusive environment for all.

When accessibility issues started to enter the academic field of audiovisual translation, scholars had different views on the concept of accessibility, which can be categorized into two main groups: a series of particularist accounts and a universalist account (Greco, 2018). For particularist accounts, accessibility only refers to certain groups of people (usually persons with disabilities) or types of barriers (for example, sensory barriers). On the other hand, the universalist account sees accessibility as an issue that concerns all human beings. Over the years, the transition from particularist to

universalist perspectives (Greco, 2018: 211) means that media accessibility no longer focuses solely on specific groups but on all people who cannot fully enjoy the media content. While more scholars are embracing the universalist account these days, this shift in media accessibility studies is not linear in time. Both particularist and universalist accounts still exist in media accessibility studies today (Greco & Jankowska, 2020).

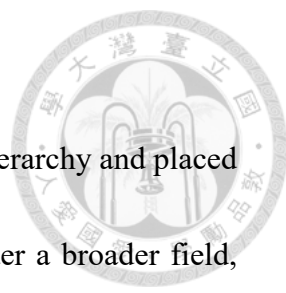
The particularist account limits media accessibility studies to Audio Description (AD) and SDH, and narrows it down to people with sensory disabilities (Greco & Jankowska, 2020: 63). As for the universalist accounts, Gambier (2003: 178) discusses that screen translation serves not only traditional audience but also audience with diverse socio-cultural backgrounds and expectations including immigrants, children, elderly people, the deaf and hard-of-hearing, and the blind and visually impaired. For example, subtitles may also serve as a tool for education purposes for young children; SDH can point out some detailed sound features and the nuanced tone of voices that are ambiguous for elderly people. Later, Díaz-Cintas (2005: 4) expanded the concept of accessibility and stated that “to lip-sync, to subtitle or to voice-over a programme” also shares the idea of accessibility. The discussions of particularist and universalist accounts show different angles to approach accessibility issues and remind us of the heterogeneity of audiences, which are valuable insights when we evaluate cases of

closed captions.



However, generally considered a research area within the field of audiovisual translation, media accessibility studies have gradually grown beyond the borders and formed a rising interdisciplinary field of its own called accessibility studies (Greco, 2018). Greco (2019) pointed out three major changes regarding how accessibility issues have evolved through time: from particularist accounts to a universalist account of accessibility; from maker-centered to user-centered approaches; and from reactive to proactive approaches. The history of accessibility has been closely related to human rights issues. The legal framework proposed by United Nations has provided foundations for discussions of linguistic rights and accessibility: the International Covenant on Civil and Political Rights (1996) reaffirmed that linguistic minorities have the right to use their own language and that this right shall not be violated or denied; the Universal Declaration of Linguistic Rights (1996) proclaimed equal rights between different linguistic groups despite differences in political or territorial statuses; the Convention of Rights of Persons with Disability (2006) highlighted the crucial link between accessibility and disability, calling upon state parties to implement policies that guarantee equal access to information, communications and other services for persons with disabilities. Under these circumstances, accessibility studies help identify barriers in a variety of groups, facilitate communications, and achieve the goal of social

inclusion and equality.



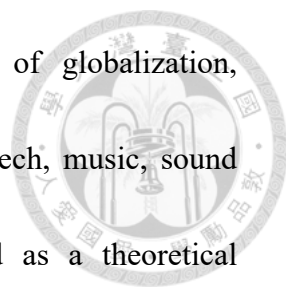
On the other hand, Orero (2021: 384) reversed the dominant hierarchy and placed audiovisual translation and translation studies as a subdomain under a broader field, media accessibility. According to Orero, as a new discipline born after the invention of a wide range of communication formats, translation studies still focus more on written texts and linguistic analysis, often leaving out other formats and languages such as sign language and music. That is why she proposed translation studies as a modality that exists under the umbrella of “accessibility.” From this perspective, translation is seen as “linguistic accessibility”, which is merely one of the many services that help overcome barriers and achieve communication for vulnerable groups (Orero, 2021: 386). In a world of rapid development of technology, the concepts and terminologies of media accessibility change and evolve through time, making it a challenging task to define and classify media services. Even so, Orero (2021: 389) tried to propose a general categorization of media accessibility and its corresponding services: to improve visual accessibility, services such as audio description, voice-over, and the new voice-command tools are used to help people with visual impairments; transcriptions and subtitles for the deaf and hard-of-hearing can help achieve audio accessibility; interactive accessibility refers to services that assist users in interacting with media content, including web accessibility and voice commands that control your smart

devices; as for linguistic accessibility, translation and interpretation are services that overcome linguistic barriers. Orero's theoretical framework and categorization showed the limit of modern translation studies that centered on a single modality and helped identify different accessibility issues as well as existing solutions.

Whether how accessibility is positioned in the research field, the fundamental idea of media accessibility is to create products, services, or environments accommodating people of all abilities and disabilities, thus facilitating communication. As the notion of "design for all" becomes more universal nowadays, intersemiotic and multimodal approaches provide a new understanding of the meaning-making process and a new framework to evaluate the quality of communication, especially in the era of new technology and new media.

2.3 Multimodality and Translation

To provide media access to all, multimodal factors must be considered during the translation process. Over the past two decades, multimodality studies emerged from a series of interdisciplinary research fields which pointed out the fact that communication involves more than just linguistic exchanges (Boria & Tomalin, 2020: 11). It challenged the dominance of monomodality (mainly written language), which has been viewed as the most esteemed form of expression (Pérez-González, 2020: 94). Due to technological



innovation and the changing media environment in the age of globalization, communication nowadays entails a variety of modes such as speech, music, sound effects, images, etc. Thus, “multimodality” has been proposed as a theoretical framework to explore the complicated meaning-making process (Boria & Tomalin, 2020: 11). Modes are defined as “semiotic resources which allow the simultaneous realisation of discourses and types of (inter)action” (Kress & Van Leeuwen, 2001: 21), such as language, image, sound, gesture, or spatial arrangement. People can combine different modes to establish meaning based on diverse communication scenarios. In addition, Kress and Van Leeuwen (2001: 20) defined multimodality as “the use of several semiotic modes in the design of a semiotic product or event”, and pointed out different ways of combining these modes. Modes may strengthen one another and express similar meanings in diverse ways or fulfill complementary roles in the meaning-making process (Kress & Van Leeuwen, 2001), such as writings and images in magazines. Pérez-González (2014: 186-187) stated that audiovisual texts involve “the simultaneous deployment of several sign repertoires”, for example, written subtitles, film editing, colors, and music. Hence, audiovisual texts are usually referred to as “multimodal”.

In Stöckl’s view, multimodality can be modeled as “a networked system of choices (as cited in Pérez-González, 2014: 191). An audiovisual text consists of specific modes

picked by the communicators to realize their communicative intentions (Stöckl, 2004: 11). To unravel the complexity of multimodality, the framework of Stöckl (2004) is chosen here to observe the modes across the visual and auditory sensory channels. He categorizes the semiotic resources in audiovisual texts into four core modes—IMAGE, LANGUAGE, SOUND, and MUSIC. Furthermore, these core modes are abstract modes that “need to be instantiated in a specific medial variant” (Stöckl, 2004: 14).

To start, IMAGE consists of two medial variants, which are static (still) images and dynamic (moving) images. While filmmakers use different types of camera movements (panning, tilting, tracking shot, etc.), editing, and visual effects to frame the moving pictures, still images are also presented in films for a variety of reasons, communicating through their color, angle, lighting, and so forth. On the other hand, LANGUAGE includes static writing, animated writing, and speech. These medial variants can be further realized through sub-modes like font, spacing, speed, volume, intonation, voice quality, and so on (Stöckl, 2004: 13). Pérez-González (2014: 194) notes that static writing (traditional intertitles and subtitles) in audiovisual texts can be used to elaborate on the actions in films or to translate the conversation of characters into another language; animated writing, in the form of moving subtitles, is now more frequently used to show the interaction between written language and the audiovisual text and convey the emotions behind words. Speech para-verbal means also help shape

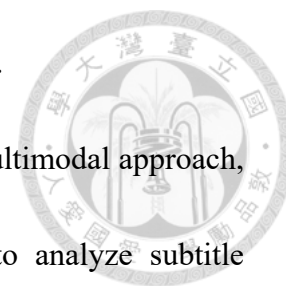
LANGUAGE in manners of hearing-related factors, such as the intonation of characters.

Modes of SOUND and MUSIC can also be manifested in the audiovisual text. In films, SOUND usually includes speech and noise (or sound effects) (Pérez-González, 2014), such as the indistinctive chattering sound in a coffee shop or the screaming sound of wounded soldiers on the battlefield. Similarly, MUSIC in films contains a number of cues that can enhance the dramatic narrative and create a rhythm that combines with the image. The core modes of IMAGE, LANGUAGE, SOUND, and MUSIC will be used in this case study to show the connections and interactions between the semiotic resources and the subtitles.

As one of the few multimodal research works on Chinese-English subtitle translation, Chuang (2006: 372) argued that subtitle translation is intersemiotic translation and that various semiotic modes in the subtitled film text contribute meanings. As a result, when we translate subtitles, meanings created by different semiotic modes should be taken into consideration (Chuang, 2006). The scholar also proposed the five semiotic modes commonly embodied in films: the written mode, the spoken mode, the mode of music, the mode of sound effects, and the mode of moving images (Chuang, 2006: 374). Compared to other subtitle translation studies, this paper indicated that subtitle translation has to deal with many semiotic modes and the translator has to combine the meanings of different semiotic modes rather than just

managing the one-on-one relationship of source text and target text.

Building on Chuang's (2006) research on subtitling from a multimodal approach, Chen (2019: 7) claimed that "multimodality is an ideal tool to analyze subtitle translation" and portrayed the process of subtitle translation based on the interactions between auditory and visual modes. Chen (2019) selected five animated movies and five live-action movies as the research data and analyzed their simplified Chinese subtitles from two major theoretical foundations: SFL (systemic functional linguistics)-informed multimodality and semiotic translation. Chen concluded the research with three major findings: Firstly, the semiotic interrelations between modes not only exist within the visual images and the synchronic subtitles of the shot but also involve visual images beyond the shot, and that is the reason why the subtitling unit is beyond the shot. It indicated that the semiotic relations in the original film might be different from the film with subtitles. Secondly, Chen suggested that the explicitation of visual images in subtitles helps ensure the audience's comprehension of the film and strengthen the narrative flow of the story, which can be a practical strategy for subtitling practices. Lastly, since the written subtitles need to combine features of the spoken expressions, Chen identified that lexis is affected more than other linguistic parts in subtitle translation. Chen's (2019) work signified the interplay between visual and auditory modes in the film text and the subtitles and pointed out how subtitles are influenced by



intersemiotic translation. To investigate the process of intersemiotic translation, Qian and Feng (2021) developed a theoretical framework to analyze Chinese-English translation.



Qian and Feng (2021) believed that Chuang's (2006) research did not provide a comprehensive theoretical framework of how subtitle translation is influenced by different semiotic modes, so they proposed a model of intersemiotic shifts as the theoretical framework, which includes Addition, Omission, Omission + Addition, Compensation, and Typographic Transformation, to illustrate and categorize the Chinese-English subtitle translation. They adopted an approach of multimodal discourse analysis to analyze the English-Chinese translation of the well-liked Chinese costume drama, *Zhenhuan Zhuan* (甄嬛傳). According to their findings, interlingual translation in general is shorter/simpler in the multimodal context, and non-verbal modes such as IMAGE and para-verbal factors play important roles in conveying meaning in the multimodal contexts (Qian & Feng, 2021: 39-40).

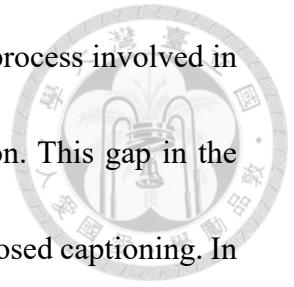
In a recent paper, Gambier (2023) discussed the changing landscape of translation studies in light of the growth of multimodal communication, driven by digital technology. The conventional Western paradigm of translation is challenged by new expectations, requirements, and practices created by the rapid development of technology, and various concepts in translation studies must be reviewed, including text,

authorship, translation unit, and quality (Gambier, 2023: 1). Among fields of translation studies, audiovisual translation is a relevant example in the new media environment, which entails a complicated system of signs (Gambier, 2023). Multimodal texts in audiovisual translation shift away from a language-centric view of translation and require new approaches to assess translation quality, which goes beyond a series of linguistic requirements (Gambier, 2023: 12).

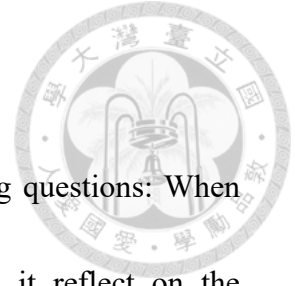
To evaluate the quality of translation, Gambier (2023) reaffirmed his position that accessibility is a significant concept in audiovisual translation and proposed three criteria to improve user-friendliness in subtitling: Acceptability, referred to as linguistic standards including styles, terminology, and rhetorical modes; Legibility, determined in regards to fonts, subtitle position, and subtitle speed; and Readability, defined by reading speed and patterns, text difficulty, and semantic load, etc. The multimodal features in today's text compelled us to reconsider the norms of translation quality established within the paradigm of equivalence and focus on the viewer experience.

Currently, there is limited research discussing the practices and government regulations of closed captioning for the deaf and hard-of-hearing community in Taiwan and other Chinese-speaking countries. While a few studies exist on media accessibility, domestic legislation, and public television captioning services (Chen, 2010; Liu & Yang, 1997), there is a notable lack of studies investigating practical captioning strategies on

streaming platforms. In addition, the multimodal meaning-making process involved in audiovisual translation has not received sufficient research attention. This gap in the literature highlights the need for more comprehensive research on closed captioning. In this research, we will adopt a holistic approach to delve into the captioning solutions in the new media landscape and shed light on the importance of media accessibility for the deaf and hard-of-hearing viewers.



3. Research Methods



The overall aim of this research is to investigate the following questions: When facing diverse target audiences with different needs, how does it reflect on the intersemiotic translation strategies? As captioners, what captioning strategies can we use to promote media accessibility for the hearing impaired? To answer these questions, we will analyze a film from a multimodal perspective. This section will introduce the selected film text, *Little Women* (2019), justify the text selection, and explain the data collection method and classification with English subtitles (closed captions) as examples for the categorization.

3.1 Film Text Selection: *Little Women* (2019)

The closed captions of the film *Little Women* (2019) are selected as the case study of this research. The film is adapted from the novel *Little Women* written by the American author Louisa May Alcott. Published in 1868, *Little Women* tells a captivating story of the March sisters—Meg, Jo, Beth, and Amy—and their journey from childhood to adulthood, drawing inspiration from the life of the author and her three sisters. As the story unfolds, the March sisters and their loving mother Marmee live in a Massachusetts town while their father is away serving as a Union Army chaplain during the American Civil War. In the March family, four sisters have different personalities

that represent diverse aspects of women. The eldest sister, Meg March, is gentle and loving, and dreams of marriage and settling into domesticity. As for the feisty and adventurous protagonist, Jo March intends to make her own way by pursuing her career as a writer in the literary world dominated by men. The third sister, Beth March, is quiet and sweet and constantly makes sacrifices for her beloved family. Lastly, the youngest Amy March is the artist of the family, who sometimes behaves in a self-centered way. The story unfolds as four sisters overcome the obstacles in life and pursue their dreams and happiness.

Little Women explores themes of female autonomy, gender roles, the necessity of work, and the importance of family and genuineness. The novel has been a popular work and translated into many languages, inspiring numerous adaptations such as films, TV series, cartoons, stage plays, radio plays, musicals, an opera, and several literary retellings by different authors. Literary critic Showalter (1991) pointed out that *Little Women* is “the American female myth”, and the heroine Jo March has become an influential figure that represents the independent and creative women in America. Therefore, the coming-of-age classic holds a special place in the hearts of many female creators, for example, Gillian Armstrong, Vanessa Caswill, Clare Niederpruem, and the writer and director Greta Gerwig.

The film *Little Women* (2019) adapted and directed by Greta Gerwig is selected for

this case study. Gerwig restructured the novel and twisted the chronology of childhood and adulthood of the March sisters, creating similar scenes in different timelines that evoke nostalgia and memory in this familiar story, which proves to be a masterstroke (Baughan, 2020).

Being the seventh film adaptation of the 1868 novel, *Little Women* (2019) highlights the connection between the author Alcott and the protagonist Jo by quoting Alcott's writing ("I'd rather be a free spinster and paddle my own canoe.") and crafting the opening and the ending scenes of Jo negotiating with the editor to publish her work, *Little Women*. The meta-narrative in the film creates an echoing effect that draws our attention to the interconnection of the author, the main character, the original work, and the adaptation. In addition, the director shows her control of mise-en-scene by carefully adapting the plots, blocking the scenes, framing the shot, and utilizing the props, to foreshadow later events, establish the relationships of characters, and compress the plots in different chapters of the novel, which shows different approaches to re-tell a story on screen.

Scored 91 out of 100 on Metascore and 7.8 out of 10 on IMDb rating, *Little Women* (2019) has received great responses from viewers and film critics. "With a stellar cast and a smart, sensitive retelling of its classic source material, Greta Gerwig's *Little Women* proves some stories truly are timeless," according to Rotten Tomatoes, the

renowned review-aggregation website. *Little Women* (2019) scores 95% Tomatometer and receives a 92% Audience Score, which shows the film is adored by both critics and audiences. In 2020, the film was nominated for six Academy Awards, including Best Picture, Best Adapted Screenplay, Best Actress, Best Supporting Actress, and Best Original Score, and won for Best Costume Design. It also earned five British Academy Film Award nominations in 2020, with a win for Best Costume Design, and two Golden Globe Award nominations.

3.2 Justification of Text Selection

Being one of the most popular adapted films in recent years, *Little Women* (2019) has achieved great box-office performance (D'Alessandro, 2020) and is streaming on major platforms such as Netflix, Apple TV, and Prime Video. As the largest media content provider in America, Netflix released the film with English subtitles (closed captions) and Chinese subtitles along with multiple-language versions, which are accessible to audiences worldwide.

To transform *Little Women* into an audiovisual work, verbal elements and non-verbal elements, like tonalities, color palettes, music, and sound effects, are employed in the film. *Little Women* (2019) uses a variety of semiotic modes in the meaning-making process, which is particularly evident in films based on literary work. When

discussing multimodality in the subtitle translation process, Chuang (2006) points out that multiple semiotic modes contribute different meanings, which shape the film text and the subtitled film text. Therefore, these semiotic modes in *Little Women* (2019) may affect the subtitling choices of English closed captions.

In *Little Women* (2019), the English closed captions are detailed transcriptions of the sound feature in the film. In order to create a better cinematic experience for the hearing impaired, closed captions consist of both speech and non-speech elements like the dialogue of characters, relevant music cues, and a variety of sound effects. In the closed captions of *Little Women* (2019), we can see word-for-word dialogue that includes all the repetitions and filler words omitted in the Chinese subtitles, such as “Ow!”, “Um...”, or “Oh.” In addition, the speaker’s tone, the sound produced by actions, and background music cues are presented in brackets, like “[WHISPERS]”, “[SIGHS]”, “[DOOR OPENS]”, or “[♪♪♪]”. The detailed-oriented subtitles are essential to convey the subtle emotions of characters, establish the background, and give hints of scene transition. For deaf and hard-of-hearing viewers, it is also important to identify who the speaker is, especially when multiple are characters talking at the same time. Therefore, we can see captions placed directly below the speakers respectively, which avoids confusion and helps the audience comprehend the dialogue more clearly.

In 2010, as one of the top online streaming platforms, Netflix started its closed-

captioning programs. In 2012, the National Association of the Deaf (NAD) in the United States filed a class-action lawsuit against Netflix, claiming that Netflix's website was a "place of public accommodation" that violated the *Americans with Disabilities Act* because Netflix failed to provide close captions on most of its "Watch Instantly" programs, which denied equal access for the deaf and hard-of-hearing community (Alvarez, 2012; National Association of the Deaf, 2012). In the end, Netflix reached a historical agreement with NAD, ensuring that the company captioned all its programs by 2014 (Mullin, 2012). According to NAD, the agreement was hailed as "a model for the streaming video industry" (Mullin, 2012). Due to its well-developed process of captioning and its established quality of subtitles, closed captions on Netflix serve as good material for this research.

In all, considering the viewership and accessibility of the film, the interactions between closed captions and the film text, the quality assured by Netflix, and Netflix's well-established captioning practices, the English subtitles (closed captions) of *Little Women* (2019) on Netflix is a suitable case to explore audiovisual translation and media accessibility. Therefore, this study focuses on the captioning of *Little Women* (2019).

Before delving into the collection and classification of captioning cases from the film, we will look into Netflix's style guide for subtitles for the deaf and hard-of-hearing, which applies to all English-language content on the platform. By understanding the

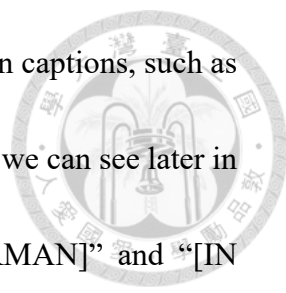
captioning guidelines established by Netflix, we can gain a deeper understanding of the basic principles of captioning on the streaming platform and relate these rules to actual captioning practices observed in *Little Women* (2019).



3.3 An Overview of Netflix's Style Guide

Netflix's style guide for the deaf and hard-of-hearing is divided into eighteen parts (see Appendix), outlining specific requirements regarding content accuracy, continuity, foreign dialogue, italics, and more. Three key features stand out in these guidelines: verbatim transcription of dialogue, prioritizing plot-relevant elements, and the use of special formatting conventions. These vital aspects ensure that closed captions accurately convey the content and effectively communicate important information to the intended audience.

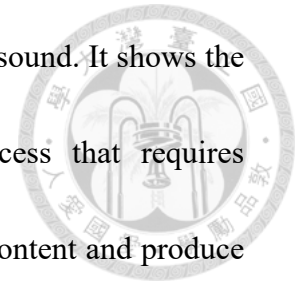
The first vital feature is the word-for-word transcription of dialogue. To provide the same access to audiovisual content as hearing audiences, captioners are asked to include as much of the original content as possible, without omission or simplification. However, when editing is necessary to match the reading speed and synchronize with the audio, text reduction and condensing are preferred over paraphrasing. Overall, it is important that dialogue should never be censored, including slang or any dialectal features. In instances when characters speak in foreign languages, dialects, or accents,



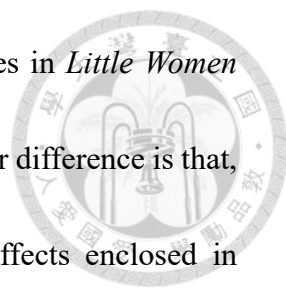
captioners must research the language being spoken and identify it in captions, such as “[speaking Italian]” rather than “[speaking a foreign language]”. As we can see later in 4.2, the captions of Case 9 clearly indicate “[SPEAKS IN GERMAN]” and “[IN ENGLISH]”, pointing out the languages for the hearing-impaired viewers. It is worth noting that not all deaf or hard-of-hearing viewers can lip-read, and the camera distance sometimes makes it difficult to observe the speaker’s mouth movement, such as in long shots that aim to show the relationship between characters and the environment. Nonetheless, the mouth movement implies the beginning and end of the sentence, and the match between captions and visual content makes it easier to follow than an edited version (Zárate, 2021: 40).

Another key aspect is prioritizing plot-relevant content. If audio interferes with dialogue, plot-pertinent content must take precedence. Ideally, captions should include all audible lines and important sound effects; nevertheless, there are space limitations and reading speed concerns in captioning practices. Captioners should opt to convey important dialogue, song lyrics, music, silences, and sound effects relevant to the narrative and convey the information in captions for the deaf and hard-of-hearing audiences. When describing sounds and music, detailed and descriptive adverbs can be used to appropriately portray the voice quality, speech speed, sound volume, and more. It is also worth mentioning that if the plot-pertinent sound effects can be inferred

through visual content, captioners can choose not to transcribe the sound. It shows the interplay of modes and the multimodal meaning-making process that requires captioners to take a holistic approach to interpret the audiovisual content and produce captions for the intended audiences.



Finally, the use of special formatting conventions plays a crucial role in promoting media accessibility, including italics, icons, and speaker identification. Italics are used in situations involving narration, inner monologue, song lyrics, dialogues heard through electronic media, scenes when the speaker is off-screen, etc. Italicizing texts reminds the deaf and hard-of-hearing viewers that the on-screen captions are not regular dialogue. Therefore, they can infer the situation based on visual clues. Another use of italics is for emphasized word in speech, which helps viewers interpret the speaker's tone and grasp the important information. Besides italics, captioners can use a music note icon to suggest that the captions are song lyrics instead of dialogue. When the speaker cannot be visually identified, speaker identification should be provided in captions. However, when a character has not yet been identified in the scene, generic terms like “[man]” or “[male voice]” are used so as not to provide information that is not yet presented in the narrative, such as the identification of the male voice in Case 2 of 4.1.



Notably, there are inconsistencies between captioning practices in *Little Women* (2019) and the rules outlined in the captioning guidelines. One major difference is that, according to the style guide, speaker identification and sound effects enclosed in brackets should be in lowercase, except for proper nouns. However, as we see in the following Results and Analysis section, the information provided in brackets is capitalized in *Little Women* (2019). This suggests that style guides may evolve over time, or captioners are afforded a degree of flexibility to better cater to audience needs and preferences regarding captioning formats. While inconsistencies exist, changes to established guidelines allow for ongoing improvement and optimization of captioning practices.

Netflix's style guide offers a reference for data analysis and provides extended details and requirements that are not covered in the scope of this research. While the analysis will not explore every aspect of the guidelines, gaining an overview of the style guide helps us understand the considerations and limitations involved in captioning. This understanding will provide valuable context before we elaborate on the collection and classification of captioning cases in *Little Women* (2019).

3.4 Data Collection and Classification

The English subtitles (closed captions) of *Little Women* (2019) on Netflix constitute

the data to be analyzed in this study. Subtitles and the corresponding screenshots of the scenes are transcribed and collected for further discussion. When analyzing the data, we will focus on the interaction of subtitles and images, and pay attention to how English subtitles (closed captions) translate sound and music for the hearing-impaired.

The intersemiotic translation strategies of closed captions of *Little Women* (2019) can be classified into three major categories proposed by Jiménez Hurtado and Martínez Martínez (2018: 121-122), which are employed here as a reference to classify the data.

To begin with, **Omission** is defined by Martí Ferriol (2013) as “the total suppression in the TT of any element in the ST” (as cited in Jiménez Hurtado & Martínez Martínez 2018, 122). It is when the sounds relevant to the storyline are omitted in the subtitles.

The second category consists of three strategies: **Categorization**, **Attribution**, and **Explanation**. Categorization refers to the cognitive activity of intersemiotic translation and consists of the conceptual category of the sound (e.g., rain), the agent producing the sound (e.g., the clerk in the shop), the action producing the sound (e.g., clapping the hands), the result of the action (e.g., laughter), and the instrument used to produce the sound (e.g., the train); Attribution means assigning qualities to a sound, which can be implicit (e.g., shouting) or explicit (e.g., somber music); Explanation is the illustration of a sound that is “not related to a specific property but rather to its relation to other entities” (Jiménez Hurtado & Martínez Martínez, 2018: 122), which may include

Categorization or Attribution, such as “orchestra resumes playing”. The third category consists of **Symbols**, **Icons**, and **Typography**, such as different positions of subtitles and the use of italics and all capital letters.



The three major categories of intersemiotic translation strategies indicate different ways of translating sound into words. However, the first category of Omission is not present in the closed captions of *Little Women* (2019). Instead, the closed captions meticulously include detailed sound elements, ranging from the intonation of characters to the music cues that help the transition of scenes in the film. Consequently, based on the classification explained earlier and the characteristics of closed captions in *Little Women* (2019), the data for multimodal analysis in this study are divided into five categories: Categorization, Attribution, Explanation, Symbols and Icons, and Typography. In each category, one to five examples are presented in the Analysis and Discussion section.

In section four, transcriptions of closed captions will be presented in the following format as seen in Table 1. From the left, the first column is the Image Number of each scene. Then, the Screenshot of the scene and closed captions are displayed in the second column. Third, a scene description is provided to give an overview of the plot, actions of characters, and the objects that appear on the screen. Lastly, closed captions are transcribed for each scene.



Table 1

Transcription of Closed Captions

Image No.	Screenshot	Scene Description	Closed Captions
1		After Jo had a fight with her sister Amy, Marmee confides in Jo that she was angry all the time when she was young.	[WHISPERS] I am not patient by nature.

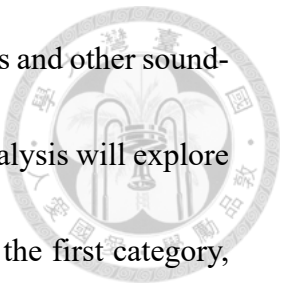
4. Results and Analysis



In this research, we will analyze how meaning is created from a multimodal perspective and examine the captioning strategies that help enhance media accessibility for the deaf and hard-of-hearing people. As mentioned in the previous section, the closed captions of *Little Women* (2019) are broken down into five categories: Categorization, Attribution, Explanation, Symbols and Icons, and Typography. In the following sub-sections, examples are provided in each category: Categorization, the most prevalent and fundamental strategy, is illustrated with five examples of labeling sound elements (see Case 1 to Case 5); Attribution, with five examples, offers additional features and details to captions (see Case 6 to Case 10); Explanation, through two examples, further elucidates the production of sounds and their relation to other entities (see Case 11 to Case 12); Symbols and Icons, with one example, employ visual representations to indicate musical cues (see Case 13); and finally, Typography is demonstrated with three examples that help identify speakers or sources of sound effects (see Case 14 to Case 16).

Throughout the following Case 1 to Case 16, we see the closed captions are all in capitals and brackets, providing messages of sound and music in written words. The capital letters of closed captions send out strong and clear signals of emotions, gestures, and ambiance that instantly bring out the feelings of the audience. As for the brackets,

it helps viewers distinguish between the verbal dialogues of characters and other sound-related cues, especially when they appear together on screen. This analysis will explore different categories and functions of closed captioning practices. In the first category, Categorization, we see how captions label the subject, action, and resulting events happening on screen.



4.1 Categorization

Categorization refers to the subject, the action, the result of the action, and the object used to generate the sound (Jiménez Hurtado & Martínez Martínez, 2018). The following cases demonstrate how closed captions put the character’s action into words and create meaning with visual factors.

Table 2

Transcription of Case 1: Categorization, exhibiting the sound resulting from the action

Image No.	Screenshot	Scene Description	Closed Captions
1		Beth insists on not telling Amy that she is seriously sick. When Jo finds out, she knows it's typical	[SIGHS]

Image No.	Screenshot	Scene Description	Closed Captions
		for Beth to do so.	

In the scene of Case 1, Jo (on the left of Image 1 of Case 1) finds out that her sister, Beth, insists on keeping her illness a secret from Amy because she doesn't want to ruin Amy's trip abroad. Upon hearing that, Jo shrugs her shoulder, takes a deep breath, and sighs resignedly, producing a slow sighing sound that expresses her weariness and sadness. Then Jo puts her arm around her older sister, Meg (on the right of Image 1 of Case 1), shakes her head with a disapproval sneer, and says "Amy has always had a talent of getting out of the hard parts of life." The scene shows Jo's feeling of sympathy and sadness for Beth's suffering, her seeking consolation from Meg, and her long-time resentfulness towards her spoiled little sister, Amy. Following Jo's reaction, Jo's mother, Marmee, gently shakes her head and reminds Jo, "Jo, don't be angry with your sister."

During this scene, the character's reaction is crucial for the audience's understanding of events on screen and the following plot. The plot-pertinent action here is enclosed with brackets as "[SIGHS]" and placed directly under the speaker. For hard-of-hearing viewers, the character's shrugging of the shoulder, the resigned and slightly irritated look on her face, and the sound effect "[SIGHS]" together formulate meanings in this scene.

In addition to actions that generate sounds, Categorization sometimes includes speakers or subjects producing the sound, helping hard-of-hearing viewers identify characters and understand the story without misconception. As a vital and distinct part of captioning practices, techniques regarding speaker identification will be discussed as follows (see Table 3).

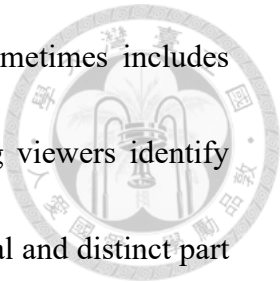


Table 3

Transcription of Case 2: Categorization, identifying speakers

Image No.	Screenshot	Scene Description	Closed Captions
1		Meg and Sallie are at the tailor shop.	MAN: Will 20 yards do?
2		Meg decides to buy silk.	Yes. MAN: Very good.

In Case 2, Meg and her friend, Sallie, are at the tailor shop. Sallie tells Meg that she should buy the beautiful silk and turn it into a new dress. After a moment of

consideration, Meg decides to buy the silk. The sales clerk asks Meg if she would like to take 20 yards, and Meg replies “Yes” with a smile. In Case 2, we can see Meg standing on the right side of the screen while talking to the clerk, who is out of sight, and on the left, Sallie appears blurry in the background (see Image 1 and Image 2 of Case 2).

The visual elements tell us that Meg is looking at and talking to someone off-screen, presumably a person standing on the audience’s right. To indicate the out-of-vision speaker, the captions are displayed as “MAN: Will 20 yards do?” and “MAN: Very good.” The “MAN:” in the captions makes the hearing-impaired audience realize that the speaker is male, not the on-screen Sallie. The male speaker here is not specified with a character’s name, but based on the dialogue and the surrounding scene of a shop, the audience can infer that he is a sales clerk. Also, the position of the captions implies the direction which the sound is coming from and thus highlights the off-camera character’s standing, creating a sense of three-dimensional space for the viewing experience.

In other cases, Categorization also indicates the instrument that produces the sound, such as the door, the train, or even the unknown object in the background. In Case 3, captions are enclosed with information on both the action and instrument (see Table 4).

Table 4

Transcription of Case 3: Categorization, depicting the action that produces the sound

Image No.	Screenshot	Scene Description	Closed Captions
1		Jo and Laurie are dancing on the porch when they hear Meg knock on window.	[KNOCKING ON WINDOW]
2		Meg gestures at them from inside, saying she is hurt.	What? My foot.

In Case 3, Jo and Meg are invited to the New Year's party, where Jo feels like an outsider and meets her like-minded friend, Laurie. Since Jo scorches her dress and cannot dance in the ballroom, Laurie suggests that they dance on the porch, away from everyone. While Jo and Laurie dance wildly outside the house, they hear a knocking sound coming from the window. It turns out that Meg hurts her foot and wishes to leave the party early. The captions "[KNOCKING ON WINDOW]" help the deaf and hard-of-hearing viewers figure out why Jo and Laurie suddenly stop dancing and the camera moves toward the window. The sound of knocking on the window might vary

depending on different factors, including the material of the window, the force of the knock, or the surrounding environment, but it is usually a short, sharp, and urgent tapping sound that attracts the attention of someone nearby. The sound creates a sense of someone trying to communicate from the other side of the window, bringing the audience's attention to the interaction between Jo and Meg.

In addition, it is worth mentioning that the caption does not specify the subject producing the sound (Meg), which is a common practice in closed captions when the subject is off-camera. Instead, the caption creates the same effect as the audiovisual language. As the audience, first, we hear the knocking sound without knowing who is making the noise, and when the characters move their way toward the window, we see Meg stand inside the house with a painful look on her face because of her injured foot. Therefore, the audience knows that it is Meg who knocks on the window earlier. By not revealing the acting subject, the caption in Image 1 of Case 3 follows the narrative flow of the film without ruining the viewing experience. The example shows that when it comes to captioning, it is not always “the more the merrier.”

As explained earlier, closed captions provide the deaf and hard-of-hearing viewers with a more comprehensive experience by indicating the subject, the action, and the instrument that generates the sound. Nonetheless, not all sound effects are clear or recognizable, which makes it challenging to transcribe the sound in detail. As we see in

Table 5, captions can be vague and general when depicting the background dialogue.



Table 5

Transcription of Case 4: Categorization, presenting background dialogue

Image No.	Screenshot	Scene Description	Closed Captions
1		After watching a play together, Jo, Meg, Laurie, and Mr. Brooke leave the theatre.	[PEOPLE CHATTERING INDISTINCTLY]

In Case 4, Jo and Meg are invited to see a play with Laurie and Mr. Brooke. After a wonderful night, they leave the theatre together. In this scene, we see the four characters walking out of the theatre amid the chattering voices of other audiences. In this extreme long shot, the theatre is the most obvious object on the screen, with characters that are too small to recognize. The exterior scene and the crowd sound effects indicate that the show is over, and the characters are heading home soon. The caption “[PEOPLE CHATTERING INDISTINCTLY]” recreates a buzzing and lively scene and helps switch from one scene to another, which keeps the audience engaged and moves the story forward. Throughout the entire film, “[PEOPLE CHATTERING

INDISTINCTLY]” appears eight times in scenes of the bustling street, the lively theatre, the upper-class ballroom, the noisy boarding house, the inaudible conversation downstairs, etc. In these scenes, the caption is not designed for understanding the dialogue but for building the atmosphere and making transitions.

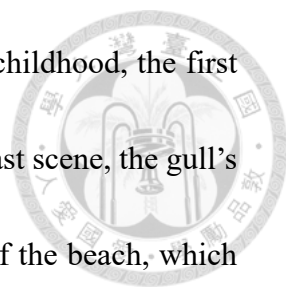
Other than human voices, captioners convey animal noises as an approach to allow the hearing-impaired audience to fully engage with the soundscape (see Table 6).

Table 6

Transcription of Case 5: Categorization, describing animal sounds

Image No.	Screenshot	Scene Description	Closed Captions
1		In the present scene, Laurie sees Amy leaving with Fred.	N/A
2		In the past scene, Laurie introduces Fred’s family to the March sisters.	[GULLS CAWING]

In Case 5, Laurie stands on the top of the staircase and sees Amy run off to greet



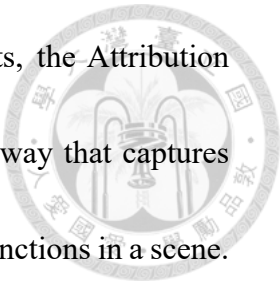
her potential husband, Fred, followed by a beach scene from their childhood, the first day Amy met Fred. During the transition between the present and past scene, the gull's cawing sound comes out slightly before the corresponding image of the beach, which is a small technique in filmmaking that creates a more seamless, natural, and immersive experience for the audience. When the audience hears the seagull's sound in Image 1 of Case 5, which might be weird at first, it allows the brain to anticipate and prepare for the following visual information, the beach. For the hard-of-hearing audience, "[GULLS CAWING]" identifies the animal and describes the loud and rough sound of the bird; however, the caption does not appear simultaneously when the gull's sound is first heard in Image 1 of Case 5. Instead, the caption is presented in synchronization with the beach scene in Image 2 of Case 5. The minor time lag of the caption destroys the cohesive and seamless flow of the scene transition, creating a clean and abrupt cut that directly takes the audience to a different time and space.

While Categorization in Case 1 to Case 5 focuses on the subject, the action, and the instrument that generates the sound effects, the Attribution strategy goes a step further by ascribing more explicit or implicit meanings to sounds.

4.2 Attribution

Attribution involves assigning qualities to sounds (Jiménez Hurtado & Martínez

Martínez, 2018: 122). Rather than merely labeling sound elements, the Attribution strategy describes and interprets the speech and sound effects in a way that captures more nuanced depths of emotions, tonal qualities, and atmospheric functions in a scene.



In Table 7, one exterior scene (Image 1 of Case 6) and one interior scene (Image 2 of Case 6) are put together to analyze the Attribution strategy.

Table 7

Transcription of Case 6: Attribution, assigning qualities to a sound

Image No.	Screenshot	Scene Description	Closed Captions
1		Jo hops over a fence and runs toward a house.	[BREATHING HEAVILY]
2		Aunt March falls asleep while Jo reads aloud to her.	[SNORING LIGHTLY]

After saying goodbye to her three sisters at the fork in the road, Jo rushes to Aunt March’s place to read aloud books to her as a way of earning some pocket money. In

Case 6, Image 1 shows that Jo jumps over a fence and runs towards Aunt March's house.

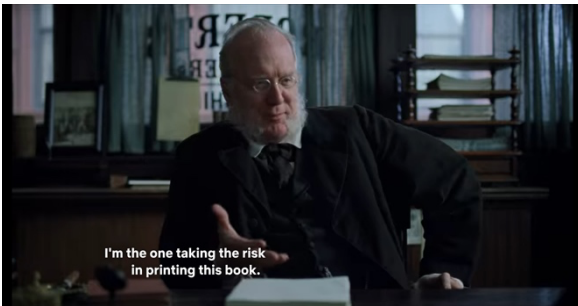
The closed captions of "[BREATHING HEAVILY]" captures Jo's huffing and puffing as she makes her way to the house, and the sound indicates that Jo is running hastily, perhaps she is running late for her errand. Meanwhile, the visual clues tell the audience that Jo chooses to take a shortcut to get to her aunt's house as fast as possible.

Subsequently, the shot changes to an interior scene of Aunt March's house, and the audience can see Aunt March sit in a chair and doze off with a poodle in her lap (see Image 2 of Case 6). While Aunt March accidentally falls asleep, the next scene shows that Jo stands next to the bookshelf and surreptitiously reads her own book, trying to be as quiet as possible. The captions of "[SNORING LIGHTLY]" here suggest Aunt March is in light sleep and might wake up at any time. The gentle snoring sound suggested by the captions allows viewers with hearing loss to recreate the scene and grasp more information happening on screen.

The descriptive adverbs in closed captions give more color to characters' actions, reflect their personalities, and help the audience understand the narrative between the scenes. In the following Table 8, the caption reflects the tone of the speaker and creates a more vivid image of the character's temperament.

Table 8

Transcription of Case 7: Attribution, describing the tone of voice

Image No.	Screenshot	Scene Description	Closed Captions
1		The editor negotiates the contract with Jo, implying that he is not willing to pay her in advance.	I'm the one taking the risk in printing this book.
2		Jo disagrees with the editor and demands a better deal.	[SCOFFS] Yes, but it's my book.

In Case 7, Jo negotiates the contract for her new book with the editor at the publishing house. The editor says that he is willing to offer 5% of the royalties, and Jo asks if she would receive an up-front payment. Upon hearing Jo's request, the editor slightly gestures his right hand and claims that "I'm the one taking the risk in printing this book", suggesting that he would not give an advance payment. Jo admits that the editor is taking the risk of publishing the book, but as the author, she believes that she deserves more than just the royalties. In the scene, the tone of Jo's voice is conveyed as "[SCOFFS]", which expresses her contempt, disapproval, and derision on this matter.

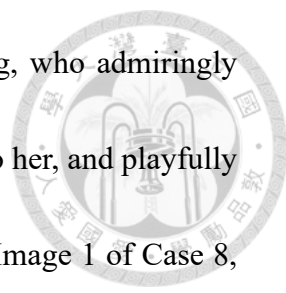
The tone of the voice can influence the meaning of the spoken words. By pointing out the speaker's tone, it helps the deaf and hard-of-hearing viewers grasp the emotions of the characters and the nuances of the conversation. The tone enclosed in brackets provides an alternative solution for conveying audio cues in dialogues, which is particularly important in sarcastic remarks or dramatic voices, such as the next example in Table 9.

Table 9

Transcription of Case 8: Attribution, describing the tone of voice

Image No.	Screenshot	Scene Description	Closed Captions
1	 A screenshot from a film showing a group of people on a beach. Laurie is in the foreground, looking towards the camera. The caption reads: [SINGSONGY] I know something you don't know.	At the beach, Laurie tries to tell Jo a secret.	[SINGSONGY] I know something you don't know.
2	 A screenshot from a film showing Jo and Laurie walking on a beach. Jo is looking at Laurie. The caption reads: Tell me this second.	Jo is intrigued and wants to know the secret.	Tell me this second.

In Case 8, the March sisters and their friends enjoy a beautiful spring day by the

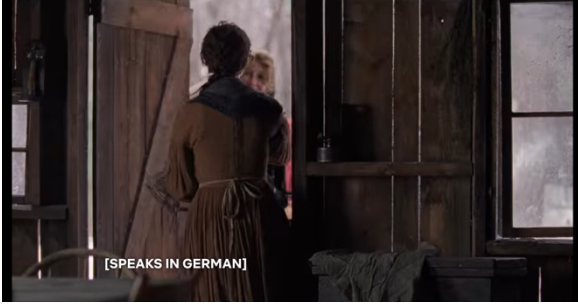


sea. While wandering on the beach, Jo keeps looking over at Meg, who admiringly gazes at Mr. Brooke. The scene annoys Jo. Laurie sees her, comes to her, and playfully mentions that he knows something about Meg and Mr. Brooke. In Image 1 of Case 8, Laurie speaks to Jo in a voice that rises and falls in levels, creating a lively rhythm as if he were singing a song. The addition of “[SINGSONGY]” in captions reflects the vibrant tone of the speaker and shows us the mischievous and playful side of Laurie’s personality. Without specifying the tone of his voice, the caption of “I know something you don’t know” can sound like a more serious and mysterious statement, which might create a different ambiance and build up a different image of Laurie.

The tone of the voice is an indispensable part of the performance that requires careful transcription. As discussed in Case 7 and Case 8, indicating the tone in captions makes the audiovisual content more comprehensible, accessible, and enjoyable for the hearing-impaired audience. Moreover, to transcribe the qualities of the sound, sometimes it is essential to point out which language the speaker uses in certain scenes (see Table 10).

Table 10

Transcription of Case 9: Attribution, identifying the languages

Image No.	Screenshot	Scene Description	Closed Captions
1		Marmee and the March girls visit the poor Mrs. Hummel.	[SPEAKS IN GERMAN]
2		Mrs. Hummel is grateful that the March family bring them food and clothes.	[IN ENGLISH] It is good angels come to us.

In Case 9, Marmee and the March girls visit the Hummel family on Christmas day. Mrs. Hummel and her five children live in a dilapidated shack near the woods, without sufficient food and clothes to get through the tough wintertime. The March family decides to give the Hummels a hand, bringing food, blankets, and medication to the poor family. After Marmee knocks on the door, Mrs. Hummel answers the door in surprise and expresses her gratitude. In Image 1 of Case 9, seeing Marmee and her daughters showing up with plenty of food and clothes that she desperately needs, Mrs. Hummel instinctively whispers in German, “Mein Gott” (My God), which shows her surprise and joy. With the gentle music score in the background, it is not easy to recognize what Mrs. Hummel says in this scene. As we can see in Image 1 of Case 9,

the caption is vaguely presented as “[SPEAKS IN GERMAN]” with no clear transcription. In the following shot, Mrs. Hummel’s two kids sit by the bed, looking weak and malnourished, and Mrs. Hummel stands by the door and says in English, “It is good angels come to us” (see Image 2 of Case 9). The change of language is indicated in the captions as “[IN ENGLISH]”, suggesting that the following conversations are now back in English.

Languages in the film can reflect the cultural background and socioeconomic status of the speaker, as well as give the audience clues about the location of the story and the time periods in history. For the hearing audience, it is relatively easy to recognize that the character is speaking in a foreign language, and the sound mode can help build up a more colorful character that has a different cultural background. In Case 9, at first, when Mrs. Hummel is surprised, she quickly whispers in German. It suggests that German could be her mother tongue, and the Hummels might be immigrants from a German-speaking country, which could explain why the family is living a tough and isolated life that requires help. If the multilanguage feature is not identified and labeled appropriately but generalized as one single language throughout the entire film, the diverse cultural elements created by the filmmakers are lost during the process of intersemiotic translation.

As a strategy of intersemiotic translation, Attribution not only involves assigning

features to the speaker’s voice but also entails giving details to music in films. We will discuss how music is transcribed in films in Table 11 that follows.



Table 11

Transcription of Case 10: Attribution, giving details to music in films

Image No.	Screenshot	Scene Description	Closed Captions
1		Meg and Mr. Brooke get married in the garden.	[BAND PLAYING UPBEAT TUNE]
2		Meg and Mr. Brooke dance happily.	[BAND PLAYING UPBEAT TUNE]

In Case 10, Meg and Mr. Brooke are married in the garden, followed by a picnic reception. After the ceremony, the guests cheer and applaud for the newlywed, and the band starts to play joyous and lively tunes to celebrate the happy moment. The music in this scene is non-lyrical and relevant to the plot. For hearing viewers, the IMAGE and the MUSIC modes work together to convey meaning for this scene, which shows

the celebration of love and family. To capture the uplifting and captivating music in words, the caption is transcribed as “[BAND PLAYING UPBEAT TUNE]” in Image 1 and Image 2 of Case 10. By giving details or descriptive words to depict the music, the viewers can comprehend the context and the purpose of music and relate to the emotions portrayed in the film. For the hearing-impaired audience, the “upbeat tune” suggests a cheerful, energetic, and usually fast-paced rhythm or melody that creates an engaging and positive atmosphere, and it also explains why the married couple and the guests start dancing blissfully in the garden. On the other hand, music can be transcribed in a more vague and abstract manner, which will be explained later in Case 13.

Compared to Categorization in intersemiotic translation, Attribution provides more details and characteristics to sound or music, which improves the viewer’s understanding and engagement with the audiovisual content. Furthermore, the Explanation strategy provides additional information on the production of sound or music that indicates the relationships between scenes (Jiménez Hurtado & Martínez Martínez, 2018: 122).

4.3 Explanation

The Explanation strategy in closed captioning contextualizes sounds by demonstrating their relationships to other elements within a scene, instead of

identifying specific properties of the sound (Jiménez Hurtado & Martínez Martínez, 2018). Explanation often builds upon the foundation of Categorization and Attribution strategies implicitly or explicitly, illustrating how sounds interconnect with the audiovisual context (Jiménez Hurtado & Martínez Martínez, 2018: 122). It allows the deaf and hard-of-hearing viewers to understand the cohesive soundscape of the scene and piece together the audio and visual elements within the narrative. To elaborate on this captioning strategy, two instances of Explanation are provided as follows.

Table 12

Transcription of Case 11: Explanation, delineating the production of a sound and its relations to other entities

Image No.	Screenshot	Scene Description	Closed Captions
1		After an awkward scene caused by Laurie, the orchestra stops for a moment.	[ORCHESTRA STOPS PLAYING]

Image No.	Screenshot	Scene Description	Closed Captions
2		People in the ballroom look at Laurie and whisper quietly.	[PEOPLE MURMURING]
3		After Laurie leaves the room, the orchestra resumes playing.	[ORCHESTRA RESUMES PLAYING]

In Case 11, Amy argues with Laurie as they walk across the ballroom. Amy is disappointed and annoyed to see Laurie become such a lazy, selfish, and miserable person after he is rejected by Jo, saying that he does nothing but dawdle away his life. Laurie is drunk and retorts sharply, “What have you done lately, oh, great artiste? Perhaps you're fantasizing about spending Fred Vaughn's fortune.” He suggests that Amy is no better than him because she tries to marry Fred Vaughn for his money, and suddenly raises his voice, “Fred Vaughn, ladies and gentlemen!” Laurie’s yelling leads to an awkward silence among the crowd, and the orchestra stops playing. For the deaf and hard-of-hearing viewers, it is difficult to notice the sudden absence of sounds, which is a powerful tool in filmmaking to build tensions, create intimacy, and establish

closeness to characters. To point out the abrupt silence, the captions here are presented as “[ORCHESTRA STOPS PLAYING]” (see Image 1 of Case 11). The example shows that captions not only document the dialogues and sound effects in films but also reflect the changes and disappearance of sounds.

The following captions also mention that people speak quietly after the brief silence, helping viewers to pick up the reactions of people in the background and immerse themselves in the atmosphere of the ballroom. Later Laurie exits the scene, and the orchestra starts playing again as if nothing ever happened. The captions of “[ORCHESTRA RESUMES PLAYING]” in Image 3 of Case 11 signify the return of the background music and everything in the ballroom seems to be back on track. The Explanation strategy used in Case 3 demonstrates the change of actions of the orchestra, and the similarity of syntax in captions of Image 1 and Image 3 of Case 11 foregrounds the differences between scenes and allows a coherent way of captioning that is friendly to viewers, making captions more readable.

Another example shows how captions provide descriptive information about a sound’s relationship with other elements rather than focusing on a specific quality. In Table 13, we see how captions depict the recurring ambient chatter in the background as a cohesive part of the overall soundscape.

Table 13



Transcription of Case 12: Explanation, delineating the production of a sound and its relations to other entities

Image No.	Screenshot	Scene Description	Closed Captions
1		Jo walks downstairs to meet Amy and Laurie.	[PEOPLE CHATTERING INDISTINCTLY]
2		Jo walks downstairs and hears voices from the living room.	[MEG LAUGHS]
3		Jo sees her family chatting in the living room.	She's a painter.
4		Jo looks at her family with a smile.	Are you excited?

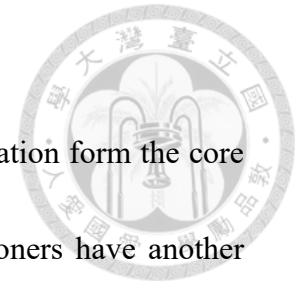
Image No.	Screenshot	Scene Description	Closed Captions
5		Jo fixes her face and prepares to greet Amy.	[CHATTERING CONTINUES INDISTINCTLY]

In Case 12, a series of screenshots are shown to illustrate the Explanation strategy.

In Image 1 of Case 12, Jo walks downstairs to meet her family. In this scene, the audience can hear a buzz of conversation, presumably coming from the floor below, and the caption is transcribed as “[PEOPLE CHATTERING INDISTINCTLY]” to indicate the unidentifiable talking voices. As Jo makes her way to the living room downstairs, we can continue to hear the laughter and the murmuring voices of the March family. The captions from Image 2 to Image 4 of Case 12 show Meg’s chuckling and the dialogue between family members, which might seem a little out of context and hard to understand for the audience. In Image 5 of Case 12, before Jo’s family sees her, Jo tries to fix her face and prepare to congratulate Amy for the engagement. The caption “[CHATTERING CONTINUES INDISTINCTLY]” helps build up the atmosphere and create a connected narrative between the shots. The “CONTINUES” in the caption indicates that the sound is not produced by a new group of people or in a different place, and it allows viewers with hearing loss to understand the meaning established by sound

effects and interpret the scenes as a whole.

While the strategies of Categorization, Attribution, and Explanation form the core conventions for conveying sounds through closed captions, captioners have another tool at their disposal: Symbols and Icons.



4.4 Symbols and Icons

The use of Symbols and Icons allows captioners to depict sounds in a more concise and vivid way. These visual representations help captioners avoid lengthy and ineffective text descriptions, transcend language barriers, and overcome the limitations of linguistic translation. However, in the film *Little Women* (2019), only one music symbol is utilized throughout the captions to represent the recurring atmospheric score. In the following Case 13, when the shot transitions from an interior scene to an exterior scene, a music symbol appears in the captions (see Table 14).

Table 14

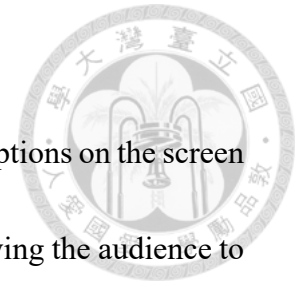
Transcription of Case 13: Symbols and Icons

Image No.	Screenshot	Scene Description	Closed Captions
1		After the scene of the four sisters in the house, we see Marmee walking alone in the snow.	[♪♪♪]

In the warm and cozy March residence, the March sisters are excited about performing a play on Christmas and passionately discuss the play together. Then the scene changes to Marmee walking towards the March house (see Case 13), and the film score appears, followed by a close shot of Marmee wiping her tears. The film score implies the change of ambiance as the audience later finds out that Marmee just finished visiting the poor Mrs. Hummel and her five children, who are almost starved to death. Since the film score is neither a song performed by a well-known artist nor a song with plot-related lyrics, the captions here are displayed as “[♪♪♪]”, offering a hint of the musical cue for the hearing-impaired. However, in other examples, the captions may be presented as “[GENTLE PIANO MUSIC PLAYING]” or “[INSTRUMENTAL MUSIC PLAYING]” to at least point out the tone, style, or specific instruments in the music. Since atmospheric music is usually instrumental and not easily recognized by the audience, it shows how challenging it can be to caption atmospherics or instrumental

music.

In addition to symbols and icons, the strategic positioning of captions on the screen plays a vital part in identifying which characters are speaking, allowing the audience to follow the film's narrative without misunderstanding or confusion.



4.5 Typography

Typography in captions encompasses choices such as font, size, letter spacing, line spacing, capitalization, italics, and bold styling. As mentioned earlier, to separate the dialogues and other sound elements, the sound effects in *Little Women* (2019) are conveyed in all capital letters enclosed within brackets. Additionally, captioners adjust the positioning of captions to enhance the hearing-impaired viewers' understanding of the storyline. By placing the captions in proximity to the corresponding characters, the deaf and hard-of-hearing viewers can intuitively match the conversation and sound effects to the characters on screen and avoid mix-ups in scenes with multiple speakers (see Case 14 and Case 15). Another approach to identifying the off-screen speaker is to italicize the captions (see Case 16), which usually happens when the character is dreaming, reminiscing, or reading letters.

Table 15

Transcription of Case 14: Typography, adjusting caption positioning

Image No.	Screenshot	Scene Description	Closed Captions
1		Amy and Aunt March sit in a carriage. Amy's mind wanders elsewhere when Aunt March calls her name.	Amy! Hm?

In the scene of Case 14, Amy and Aunt March sit side by side in a carriage (see Table 15). Aunt March talks about how French women these days are too weak to lift a hairbrush, but Amy is reading the letter from her family without noticing her aunt is talking to her. When Aunt March finds that Amy doesn't respond, she calls Amy's name to draw her attention. In Image 1 of Case 14, we can see that the positions of the captions also reflect two speakers' positions in this scene. On the right is Aunt March calling Amy's name, and on the left is Amy replying to her call. When captions are placed beneath the characters respectively, the hearing-impaired viewers can instantly recognize the speakers, which helps them understand the story better. Without this information, the deaf and hard-of-hearing viewers might mistake one character's saying for another's, especially when multiple characters talk simultaneously in one scene. It

creates obstacles to understanding the storyline and might result in misinterpretation.

When multiple sounds are produced at the same time, the position of the captions makes it easier to identify the characters who generate the sound. As shown below, the captions in the following case demonstrate how to transcribe sounds from various characters (see Table 16).

Table 16

Transcription of Case 15: Typography, adjusting caption positioning

Image No.	Screenshot	Scene Description	Closed Captions
1		After Jo says goodbye to Friedrich, she turns around to find everyone staring at her.	[STIFLING LAUGH] [CLEARS THROAT]

In Case 15, Jo bids Friedrich farewell and closes the door, and she turns back to find everyone staring at her. Jo’s family is curious about the gentleman and assumes that he and Jo have a romantic relationship. In Image 1 of Case 15, Mr. March (on the right) clears his throat to attract Jo’s attention, expressing his opinions about Friedrich and his wish to become friends with him. The noise made by Mr. March is transcribed as “[CLEARS THROAT]”, which is placed on the right to indicate that the sound comes

from the March couple. On the left, Amy covers her mouth with both of her hands, trying to hold back a laugh, and the caption here is presented as “[STIFLING LAUGH]”. These captions help the audience identify the source of the sound and interpret the emotions of the characters.

When multiple characters appear together on screen or various sounds are produced simultaneously, captions must help the hearing-impaired viewers identify the speakers. To do this, captioners generally provide speaker IDs or modify the position of the captions to match the speakers. However, the position of the captions has its limitations. Much like conventional translated subtitles, captions should avoid obscuring important visual elements on the screen. In addition to rearranging the position of captions, italics can be used to indicate the off-screen speaker (see Table 17).

Table 17

Transcription of Case 16: Typography, italicizing captions

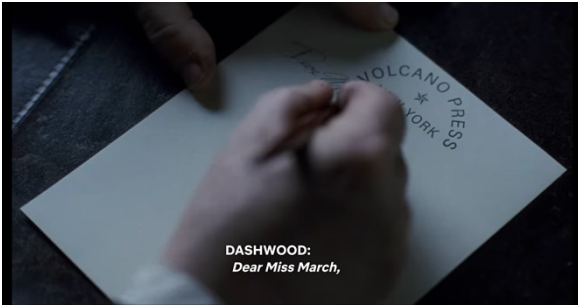
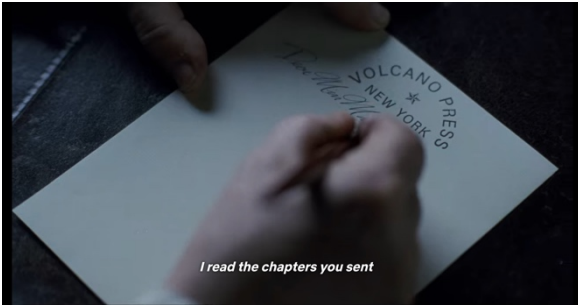
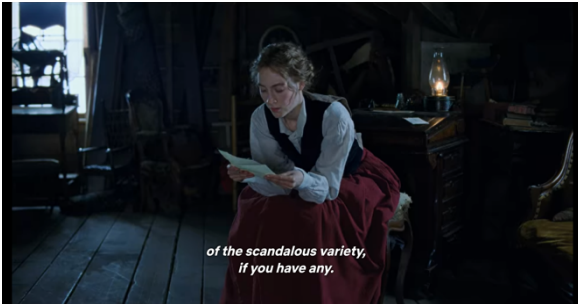
Image No.	Screenshot	Scene Description	Closed Captions
1		Dashwood writes a letter to Jo.	DASHWOO D: <i>Dear Miss March,</i>

Image No.	Screenshot	Scene Description	Closed Captions
2		Dashwood writes a letter to Jo.	<i>I read the chapters you sent</i>
3		Dashwood looks directly into the camera and expresses his rejection of the story.	and I have to agree, they aren't very promising.
4		Jo reads the letter in the attic.	<i>But, please, send more stories</i>
5		Jo reads the letter in the attic.	<i>of the scandalous variety, if you have any.</i>

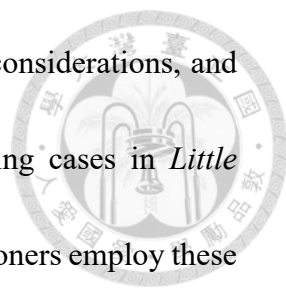
In Case 16, after the editor, Mr. Dashwood, receives the draft from Jo, he starts to write a letter to her. It turns out that Mr. Dashwood is not interested in publishing Jo's story and suggests she write something more scandalous that caters to the market.

During the scenes, Mr. Dashwood reads aloud the letter and even speaks directly to the camera (see Image 3 of Case 16), and at the same time, we see Jo read the letter in the attic. In Image 1, Image 2, Image 4, and Image 5 of Case 16, the captions are italicized to indicate the voice-over reading off the screen, and the speaker's name is identified in Image 1 of Case 16 to avoid confusion.

Even though we can see Mr. Dashwood's hands on screen in Image 1 and Image 2 of Case 16, the captioner chooses to point out the speaker to enhance comprehension for hearing-impaired viewers. In addition, italics are essential in this scenario. In the following scenes of Jo reading in the attic, hearing audience can easily realize that the male voice belongs to Mr. Dashwood. Without the use of italics, it is possible for the deaf and hard-of-hearing viewers to misinterpret the captions in Image 4 and Image 5 of Case 16 as Jo's speaking since it is natural to assume that the speaker is on screen. This case highlights the importance of utilizing italics in captions to indicate the off-screen speaker.

4.6 Discussion and Implications

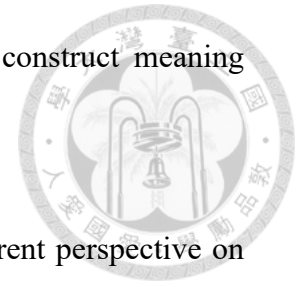
Closed captions serve as an essential assistive tool for audiovisual translation, enabling deaf and hard-of-hearing viewers to access information (Pérez-González, 2014). To achieve this goal, captioners must carefully navigate three considerations



during the captioning process: linguistic considerations, technical considerations, and special requirements (Zárate, 2021). The analysis of the captioning cases in *Little Women* (2019) and Netflix's style guide offer insights into how captioners employ these considerations. From a linguistic perspective, transcribing the original content as accurately as possible remains important, including accents, slang, and different dialectal features. Only in rare instances, when reading speed or audio-visual synchronicity poses a challenge, is text editing considered a viable option. Typographical arrangements, such as italicizing off-screen speech and enclosing sound effects in brackets, are incorporated into the captioning practice as well. Regarding technical considerations, captioners must adhere to the maximum reading speed of 20 characters per second for adult programs as stipulated by the style guide. Synchronization between captions and sound/image is another vital factor. In Case 5, when a lag occurs between the sound effect and the corresponding image, captioners must carefully time the caption to avoid confusing the audience.

Lastly, special requirements involve speaker identification, paralinguistic cues, music, sound effects, etc. As exemplified in Case 7, the tone of the voice indicated in the caption conveys nuanced meaning and vividly captures the speaker's emotions. Sound effects and music are also presented in captions, such as the background dialogue in Case 4, the orchestra's music and abrupt silence in Case 11, and the ambient music

cue in Case 13, allowing viewers with hearing impairments to construct meaning together with the visual information.



The discussion surrounding media accessibility offers a different perspective on audiovisual translation as a means to achieve social inclusion, equal access, and participation for all in this digital age. Greco (2019) identified three significant shifts in accessibility issues over time. To begin with, the perspective has gradually evolved from a particularist view of accessibility to a universalist account, recognizing that accessibility concerns extend beyond persons with disabilities to encompass all human beings, including children, seniors, migrants, and more. From a universalist perspective, Gambier (2003) pointed out that screen translation can benefit a wide range of audiences from diverse socio-cultural backgrounds. Closed captions, traditionally designed for the deaf and hard-of-hearing, can benefit a diverse range of individuals, such as language learners, those experiencing age-related hearing loss, and people enjoying the audiovisual content in noisy or quiet environments where audio is challenging (Gambier, 2003: 178). Captions extend beyond dialogue transcription, providing important details like tone of voice, sound effects, and speaker identification. These elements allow viewers to better comprehend the plot and immerse themselves in the narrative. The captioning cases in *Little Women* (2019) exemplify this, offering

information about the scornful tone employed by a speaker in Case 7, the cawing sound effect of a seagull in Case 5, and the identification of a male clerk's voice in Case 2.

Secondly, accessibility services have shifted from a maker-centered to a user-centered design approach (Greco, 2019; Romero-Fresco, 2021). Unlike open captions burned into videos, closed captions offer flexibility, enabling users to turn on or turn off the captions as needed. This user-initiated function on streaming platforms ensures universal access to this assistive form of audiovisual translation whenever desired. For example, Netflix provides captions, subtitles, and audio descriptions that can be easily switched on or off with just a click, making these accessibility services available across smartphones, tablets, TVs, and web browsers.

Last but not least, accessibility services have evolved from a reactive to a proactive approach. Once sued for lack of captioning services, Netflix now proactively provides closed captions in many programs and allows users to customize caption size, colors, font, and background colors for optimal viewing experience. Additionally, as discussed in 3.3, Netflix has established a style guide for subtitles for the deaf and hard-of-hearing, outlining general rules for transcribing dialogue in detail, prioritizing plot-relevant elements, and utilizing special formatting to improve accessibility. The analyzed cases in *Little Women* (2019) also demonstrate how captioners translate important sound effects and music into words (see Case 1, 3, 4, 5, 6, 10, 13), adjust the position of

captions to reflect the speakers (see Case 14), and indicate when the plot-relevant music abruptly ends (see Case 11). These proactive captioning options, guidelines, and strategies ensure a wide viewership for audiovisual content, creating a mutually beneficial situation for both viewers and the company.

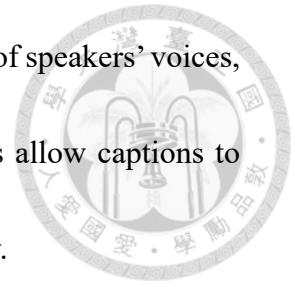
While the concept of media accessibility highlights the goal of making audiovisual translation accessible to all, the multimodality framework unlocks the complex process of meaning construction and identifies meanings created by different semiotic resources that are lost for those with sensory disabilities, thus requiring assistive approaches to recreate the meanings. The multimodal nature of audiovisual content challenges the language-centric view of translation and inspires scholars to study audiovisual translation through a more comprehensive framework (Boria & Tomalin, 2020). As Pérez-González (2014) stated, audiovisual texts entail the employment of varied semiotic resources. The meaning-making process in audiovisual work involves not only linguistic elements but also the interwoven relationships between semiotic modes, including visual, audio, spatial, gestural modes, and more.

Chuang (2006) highlighted the importance of recognizing that meaning is constructed through various semiotic modes during the subtitle translation process. Among the semiotic resources in audiovisual texts, Stöckl (2004) proposed four core modes: IMAGE, LANGUAGE, SOUND, and MUSIC. As demonstrated in the Results

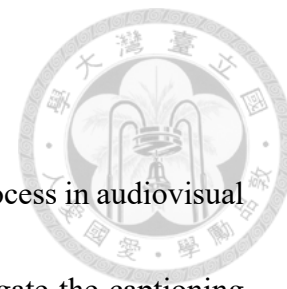
and Analysis section, these modes collectively formulate meaning, and captioners employ intersemiotic translation strategies to improve accessibility for the hearing-impaired. For instance, in Case 3, the caption conveying the action of knocking on a window (LANGUAGE), the camera movement showing the subject performing the action (IMAGE), and the abrupt knocking sound (SOUND) all contribute to meaning-making in this scene. For viewers with hearing loss, the LANGUAGE mode compensates for the lack of SOUND mode, providing plot-pertinent elements for better comprehension. Additionally, Case 9 exemplifies how captions (LANGUAGE) provide speech-related details (SOUND), indicating the spoken languages and adding depth to the character and the story background.

The interplay of these modes demonstrates the complex meaning-making process in audiovisual work and underscores the importance of utilizing intersemiotic translation strategies in captioning practice. Captioners employ intersemiotic translation strategies of Categorization, Attribution, Explanation, Symbols and Icons, and Typography to re-establish meaning potentially lost due to the modes of SOUND and MUSIC being inaccessible to hearing-impaired viewers (Jiménez Hurtado & Martínez Martínez, 2018). The Categorization strategy, for instance, provides details about the subject producing the sound, the action performed, the consequence of that action, and any object used to generate the sound. Furthermore, the Attribution strategy

aims to recreate crucial details such as the quality, volume, and tone of speakers' voices, as well as indicate the specific spoken languages. These strategies allow captions to fully capture the multimodal meaning to ensure viewer accessibility.



5. Conclusion



The study aims to explore the multimodal meaning-making process in audiovisual work, examine the intersemiotic translation strategies, and investigate the captioning practices that promote media accessibility for the deaf and hard-of-hearing community. Through the analysis of closed captions in *Little Women* (2019) and Netflix's style guides, it is observed that semiotic modes, including language, image, music, and sound collectively contribute to establishing meaning in audiovisual content (Stöckl, 2004). To compensate for the loss of sound and music modalities, captions convey speech, paralinguistic features, and plot-relevant sound messages for viewers experiencing hearing loss. To cater to their needs, captioners transcribe dialogue verbatim and adopt intersemiotic translation strategies such as categorization, attribution, explanation, symbols and icons, and typography to provide detailed information (Jiménez Hurtado & Martínez Martínez, 2018).

Captioning audiovisual content demands a multitude of techniques and expertise. The ultimate goal of providing captioning services is to allow audiences to fully appreciate the work, even in the absence of the auditory modality. It has become crucial to move beyond the language-centric viewpoint in audiovisual translation. Instead, the scholarly focus must shift towards a comprehensive understanding of the multimodal processes involved in constructing meaning across various semiotic modes, just as

Pérez-González (2014) pointed out the growing need to study the interdependence of semiotic modes in audiovisual works as these works continue to involve increasingly complex combinations of semiotic resources. Moreover, ongoing efforts should be made to reflect on approaches that can create inclusive communication experiences with maximum accessibility. In addition to closed captions, other assistive forms of audio translation such as audio description, sign language interpretation, augmented reality, and virtual reality applications have the potential to improve media accessibility and create immersive experiences tailored to the needs of diverse viewers.

While this study provides insights into the multimodal meaning-making process in captioning practices and the intersemiotic translation strategies employed by captioners to promote media accessibility, certain limitations must be acknowledged. To begin with, this research focuses on a single audiovisual work, the film *Little Women* (2019), and the captioning guidelines of a particular online streaming platform, Netflix. Future studies could benefit from exploring different genres of audiovisual content across a variety of channels to investigate potential variations in captioning strategies and practices.

Furthermore, this research primarily adopts a textual analysis methodology, drawing inferences from captioning cases within the audiovisual work. Conducting interviews with captioners could shed light on the captioning strategies employed,

challenges encountered, and decision-making processes, thereby providing valuable insights into accessibility services. Additionally, combining user studies that involve viewers with different degrees of hearing impairment would offer diverse perspectives and contribute to optimizing captioning services.

As this study has demonstrated, closed captions play a vital role in promoting media accessibility and fostering inclusive experiences for viewers with hearing loss. The author hereby calls for increased attention and efforts to provide comprehensive captioning services in Taiwan. Official statistics from the Ministry of Health and Welfare (2021) indicated that Taiwan's population with hearing impairment reached 124,485 in 2019, reflecting a 7.9% increase since 2009. In addition, as our country grapples with an aging population, age-related hearing loss is poised to become a prevalent issue. As a result, a growing number of individuals are likely to rely on closed captioning services to fully enjoy audiovisual content in the future.

Despite the growing need, hearing loss is often regarded as an invisible disability, which can be easily overlooked in our society. Therefore, it is recommended that the Taiwanese government establish comprehensive regulations requiring the provision of captioning services for TV programs, streaming platforms, and public media content. Such legislation would ensure equal access for the deaf and hard-of-hearing people. Moreover, these legislations should be regularly reviewed and updated to reflect

technological advancements and evolving user needs.

It is hoped that through legislative measures, Mandarin closed captioning services can be widely accessible across different media platforms in the foreseeable future. To provide high-quality Mandarin captioning services, captioners can employ strategies of categorization, attribution, explanation, symbols and icons, and typography as foundational strategies, while developing localized captioning approaches tailored to Mandarin-speaking viewers with hearing impairments. In addition, government authorities or captioning agencies could organize training workshops to introduce comprehensive knowledge of captioning services and practical techniques from experienced professionals. Most importantly, the development of closed captions should always involve the participation of deaf and hard-of-hearing audiences. It is crucial to design “with” them rather than design “for” them, ensuring that their needs are fully accommodated.

The objective is to create an environment where hearing-impaired individuals can seamlessly enjoy audiovisual works from all over the world, facilitated by widespread access to both intralingual and interlingual closed captioning solutions. By adopting a proactive stance towards offering assistive services and ensuring media accessibility for all members of society, Taiwan can create an inclusive society that empowers and embraces all individuals.

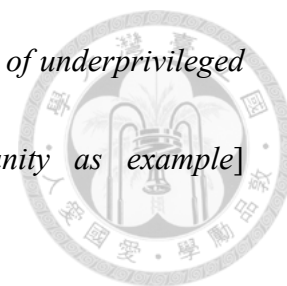
References



- Alvarez, E. (2012, October 11). Netflix to bring closed captioning to all video content by 2014. *Engadget*. <https://www.engadget.com/2012-10-11-netflix-closed-captioning-2014.html>
- Baughan, N. (2020, December 11). Little Women emancipates Louisa May Alcott's spirited sisters. *Sight and Sound*. <https://www.bfi.org.uk/sight-and-sound/reviews/little-women-2019-greta-gerwig-saoirse-ronan-florence-pugh-emma-watson-eliza-scanlen>
- Bogucki, Ł. (2019). *Areas and Methods of Audiovisual Translation Research* (3rd ed.). Peter Lang.
- Bogucki, Ł., & Deckert, M. (2020). *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (1st ed.). Palgrave Macmillan.
- Bordwell, D. & Thompson, K. (2010). *Film Art: An Introduction*. McGraw-Hill.
- Boria, M., & Tomalin, M. (2020). Introduction. In M. Boria, Á. Carreres, M. Noriega-Sánchez, & M. Tomalin (Eds.), *Translation and Multimodality: Beyond Words* (1st ed., pp. 1-23). Routledge.
- British Broadcasting Corporation. (2024, January). *BBC Subtitle Guidelines*. <https://www.bbc.co.uk/accessibility/forproducts/guides/subtitles/>
- Chen, H-W. (2010). 數位電視平台與弱勢團體媒體近用：以公共電視台服務聽障

社群為例 [Digital TV platform and the right of media access of underprivileged group: Take PTS service for hearing impaired community as example]

[Unpublished master's thesis]. National Chengchi University.



Chen, Y. (2019). *Translating Film Subtitles into Chinese: A Multimodal Study*. Springer.

Chuang, Y-T. (2006). Studying Subtitle Translation from a Multi-modal Approach.

Babel, 52(4), 372–383. <https://doi.org/10.1075/babel.52.4.06chu>

D'Alessandro, A. (2020, April 7). 'Little Women,' Big Profits: Remake Lands at No. 24

In Deadline's 2019 Most Valuable Blockbuster Tournament. *Deadline*.

[https://deadline.com/2020/04/little-women-movie-profit-2019-greta-gerwig-](https://deadline.com/2020/04/little-women-movie-profit-2019-greta-gerwig-sony-1202901744/)

[sony-1202901744/](https://deadline.com/2020/04/little-women-movie-profit-2019-greta-gerwig-sony-1202901744/)

Described and Captioned Media Program. (n.d.). *Public Law 85-905*.

<https://dcmp.org/learn/212-public-law-85-905>

Diaz Cintas, J. (2020). The Name and Nature of Subtitling. In Ł. Bogucki & M. Deckert

(Eds.), *The Palgrave Handbook of Audiovisual Translation and Media*

Accessibility (1st ed., pp. 149-171). Palgrave Macmillan.

Diaz Cintas, J., & Anderman, G. (2009). *Audiovisual Translation: Language Transfer*

on Screen. Palgrave Macmillan.

Diaz Cintas, J., & Remael, A. (2007). *Audiovisual Translation: Subtitling*. Routledge.

Federal Communications Commission. (2021, January 27). *Closed Captioning on*

Television. <https://www.fcc.gov/consumers/guides/closed-captioning-television>

Gambier, Y. (2003). *Screen Translation: Special Issue of The Translator*. Routledge.

Gambier, Y. (2012). The position of audiovisual translation studies. In C. Millan & F.

Bartrina (Eds.), *The Routledge Handbook of Translation Studies*, (pp. 45–59).

Routledge.

Gambier, Y. (2023). Audiovisual Translation and Multimodality: What Future? *Media*

and Intercultural Communication: A Multidisciplinary Journal, 1(1), 1-16.

<https://doi.org/10.22034/mic.2023.167451>

Greco, G. M. (2018). The Nature of Accessibility Studies. *Journal of Audiovisual*

Translation, 1(1), 205–232. <https://doi.org/10.47476/jat.v1i1.51>

Greco, G. M. (2019). The Rise of Accessibility Studies: Abuses, Misuses and the

Method of Poietic Design. In C. Stephanidis (Eds.), *HCI International 2019-Late*

Breaking Papers, (pp. 15-27). Springer.

Greco, G. M., & Jankowska, A. (2020). Media Accessibility Within and Beyond

Audiovisual Translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave*

Handbook of Audiovisual Translation and Media Accessibility (1st ed., pp. 57-81).

Palgrave Macmillan.

Hurtado, C. J., & Gallego, S. M. (2013). Multimodality, translation and accessibility: a

corpus-based study of audio description. *Perspectives*, 21(4), 577–594.

<https://doi.org/10.1080/0907676x.2013.831921>



Jakobson, R. (2000). On Linguistic Aspects of Translation. In L. Venuti (Eds.),

Translation Studies Reader (pp. 113–118). Routledge.

Jankowska, A. (2020). Audiovisual media accessibility. In E. Angelone, M.

Ehrensberger-Do, & G. Massey (Eds.), *The Bloomsbury Companion to Language*

Industry Studies (pp. 231–260). Bloomsbury Academic.

<http://dx.doi.org/10.5040/9781350024960.0015>

Jiménez Hurtado, C., & Martínez Martínez, S. (2018). Concept Selection and

Translation Strategy: Subtitling for the Deaf and Hard-of-Hearing. *Linguistica*

Antverpiensia, New Series: Themes in Translation Studies, 14, 114–139.

Kress, G., & Van Leeuwen, T. (2001). *Multimodal Discourse: The Modes and Media of*

Contemporary Communication. Oxford University Press.

Kress, G. (2020). Transposing Meaning: Translation in a Multimodal Semiotic

Landscape. In M. Boria, Á. Carreres, M. Noriega-Sánchez, & M. Tomalin (Eds.),

Translation and Multimodality: Beyond Words (1st ed., pp. 24-48). Routledge.

Liu, Y-L., & Yang, C-C. (1997). 傳播科技的另一種選擇——我國隱藏式字幕的政

策研究 [Another Choice for Communications Technology: An Analysis of Policy

Implications for Closed-Captioning Services in Taiwan]. *廣播與電視*, 3(2), 109-

140.

Ministry of Health and Welfare, Republic of China (Taiwan). (2021). 世界聽力

日衛生福利統計通報 [Ministry of Health and Welfare World Hearing Day

Statistics Report]. <https://dep.mohw.gov.tw/DOS/cp-5112-63001-113.html>

Mullin, J. (2012, October 11). Netflix agrees to subtitle all films by 2014. *CNN*.

[https://edition.cnn.com/2012/10/11/tech/web/netflix-subtitles-](https://edition.cnn.com/2012/10/11/tech/web/netflix-subtitles-settlement/index.html)

[settlement/index.html](https://edition.cnn.com/2012/10/11/tech/web/netflix-subtitles-settlement/index.html)

National Association of the Deaf. (2012, June 19). *Landmark Precedent in NAD vs.*

Netflix. <https://www.nad.org/2012/06/19/landmark-precedent-in-nad-vs-netflix/>

National Association of the Deaf. (n.d.). *Television Decoder Circuitry Act*.

<https://www.nad.org/resources/civil-rights-laws/television-decoder-circuitry-act/>

National Institute of Standards and Technology. (2016, September 21). *Closed*

Captioning for the Hearing Impaired: How it Originated.

[https://www.nist.gov/pml/time-and-frequency-division/time-services/closed-](https://www.nist.gov/pml/time-and-frequency-division/time-services/closed-captioning-hearing-impaired-how-it-originated)

[captioning-hearing-impaired-how-it-originated](https://www.nist.gov/pml/time-and-frequency-division/time-services/closed-captioning-hearing-impaired-how-it-originated)

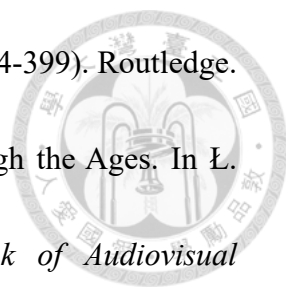
Neves, J. (2008). Ten Fallacies about Subtitling for the d/Deaf and the Hard of Hearing.

Journal of Specialised Translation, 10, 128–143.

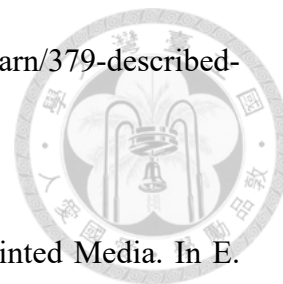
Neves, J. (2005). *Audiovisual Translation: Subtitling for the Deaf and Hard-of-Hearing*

[Unpublished doctoral dissertation]. University of Roehampton.

Orero, P. (2022). Translation, accessibility and minorities. In E. Bielsa (Eds.), *The*

- 
- Routledge Handbook of Translation and Media* (1st ed., pp. 384-399). Routledge.
- Perego, E., & Pacinotti, R. (2020). Audiovisual Translation through the Ages. In L. Bogucki, & M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and Media Accessibility* (1st ed., pp. 33-56). Palgrave Macmillan.
- Pérez-González, L. (2014). *Audiovisual Translation: Theories, Methods and Issues*. Routledge.
- Pérez-González, L. (2020). From the “Cinema of Attractions” to *Danmu*: A Multimodal-theory Analysis of Changing Subtitling Aesthetics across Media Cultures. In M. Boria, Á. Carreres, M. Noriega-Sánchez, & M. Tomalin (Eds.), *Translation and Multimodality: Beyond Words* (1st ed., pp. 94-116). Routledge.
- Qian, H., & Feng, D. (2021). Intersemiotic Shifts in the Translation of Chinese Costume Drama Subtitles. In M. Zhang, & D. Feng (Eds.), *Multimodal Approaches to Chinese-English Translation and Interpreting* (pp. 16-42). Routledge.
- Romero-Fresco, P. (2021). Creativity in media accessibility: A political issue. *The Journal of Intercultural Mediation and Communication*, 14, 162-197.
- Showalter, E. (1991). *Sister's Choice: Tradition and Change in American Women's Writing. The Clarendon Lectures 1989*. Oxford University Press.
- <https://doi.org/10.1093/acprof:oso/9780198123835.001.0001>
- Stark, B. *Described and Captioned Media Program: Seventy Years of Progress*.

Described and Captioned Media Program. <https://dcmp.org/learn/379-described-and-captioned-media-program-seventy-years-of-progress>



Stöckl, H. (2004). In *Between Modes: Language and Image in Printed Media*. In E.

Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on Multimodality* (pp.

9-30). John Benjamins Publishing Company. <https://doi.org/10.1075/ddcs.6.03sto>

Szarkowska, A., Krejtz, I., Klyszejko, Z., & Wieczorek, A. (2011). Verbatim, Standard,

or Edited? Reading Patterns of Different Captioning Styles Among Deaf, Hard of

Hearing, and Hearing Viewers. *American Annals of the Deaf*, 156(4), 363–378.

<https://www.jstor.org/stable/pdf/26235165.pdf>

Szarkowska, A. (2013). Towards Interlingual Subtitling for the Deaf and the Hard of

Hearing. *Perspectives* 21 (1), 68–81.

Szarkowska, A. (2020). Subtitling for the Deaf and the Hard of Hearing. In Ł. Bogucki,

& M. Deckert (Eds.), *The Palgrave Handbook of Audiovisual Translation and*

Media Accessibility (1st ed., pp. 249-268). Palgrave Macmillan.

United Nations. (2006, December 12). *Convention on the Rights of Persons with*

Disabilities.

[https://www.ohchr.org/en/instruments-](https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities)

[mechanisms/instruments/convention-rights-persons-disabilities](https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities)

U.S. Department of Justice Civil Rights Division. (n.d.). *Americans with Disabilities*

Act of 1990, As Amended. <https://www.ada.gov/law-and-regs/ada/>

Whyatt, B. (2017). Intralingual Translation. In J.W. Schwieter & A. Ferreira (Eds.), *The Handbook of Translation and Cognition* (pp. 1776-192). John Wiley & Sons.

<https://doi.org/10.1002/9781119241485.ch10>



Wu, C-Y. (2015). 看或不看？由你自己決定！淺談數位無線電視之隱藏字幕及聽障字幕 [To Watch or not to Watch? It's Up to You. An Overview of Closed Captions and Subtitles for Hard Hearing in Digital Terrestrial Television]. *NCC NEWS*, 8(9), 17-18.

Zárate, S. (2021). *Captioning and Subtitling for d/Deaf and Hard of Hearing Audiences*.

UCL Press. <https://doi.org/10.14324/111.9781787357105>

Zhang, M., & Feng, D. (2021). Introduction: Multimodal Approaches to Chinese-English Translation and Interpreting. In M. Zhang & D. Feng (Eds.), *Multimodal Approaches to Chinese-English Translation and Interpreting* (pp. 1-15). Routledge.

Appendix: Netflix's Style Guide for Subtitles for the Deaf and Hard of Hearing



This section applies to subtitles for the deaf and hard of hearing created for English language content (i.e. intralingual subtitles).

I.1. Accuracy of content

- Include as much of the original content as possible.
- Do not simplify or water down the original dialogue.
- Where content has been dubbed into English, please refer to the dubbing script or dubbed audio as the basis for the SDH file and ensure that the two match as much as reading speed and timings allow.
- Truncating the original dialogue should be limited to instances where reading speed and synchronicity to the audio are an issue.
- When editing for reading speed, favor text reduction, deletion and condensing but do not paraphrase.
- Transcription of the source language should follow the word choice and sentence order of the spoken dialect. Slang and other dialectal features should not be changed.

I.2. Character Limitation

- 42 characters per line.



I.3. Continuity

- When including ellipses in subtitles, please use the single smart character (U+2026) as opposed to three dots/periods in a row.

- Do not use ellipses or dashes when an ongoing sentence is split between two or more continuous subtitles (e.g. when the pause between utterances is less than 2 seconds)

Subtitle 1 I always knew

Subtitle 2 that you would eventually agree with me.

- Use an ellipsis to indicate a pause (2 seconds or more) or if dialogue trails off. In the case of a pause of under two seconds, if the sentence continues in the next subtitle, do not use an ellipsis at the beginning of the second subtitle.

Subtitle 1 Had I known...

Subtitle 2 I wouldn't have called you.

- Use two hyphens to indicate abrupt interruptions.

-What are you--

-Be quiet!

-What are you--

-[bomb explodes]

- Use an ellipsis followed by a space when there is a significant pause or hesitation *within* a subtitle.

She hesitated... about accepting the job.

- Use an ellipsis without a space at the start of a subtitle to indicate that a subtitle is starting mid-sentence.

...have signed an agreement.

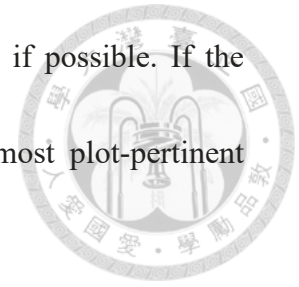
I.4 Dates and Decades

- Dates should always be written in the order in which they are said (i.e. as per the audio) but omitting words like "the" and "of", i.e. 6th March or March 6th, not the 6th of March.
- Decades should be written using numerals in the following format: nineteen fifties should be 1950s, fifties should be '50s.
- Centuries should be written in the following format: twentieth century should be 20th century.
- Do not use '50s, '70s etc. for ages: i.e. prefer "I am in my fifties" vs. "I am in my '50s" or "I am in my 50s".

I.5. Documentary/Unscripted



- For TV/movie clips, all audible lines should be transcribed, if possible. If the audio interferes with dialogue, please give precedence to most plot-pertinent content.
- Avoid going back and forth between italicized and non-italicized subtitles when the speaker is on and off screen. If the speaker is on-camera for at least part of the scene, do not italicize. Leave italics for off-screen narrators.



I.6. Dual Speakers/Multiple Events

- Use a hyphen without a space to indicate two speakers in one subtitle, with a maximum of one speaker per line.

-Are you coming?

-In a minute.

- When identifiers are needed, they should follow the hyphen as follows:

-[Kimmy] Are you coming?

-[Titus] In a minute.

-[Kimmy] Are you coming?

-In a minute.



- Hyphens are also used to indicate a speaker and a sound effect, if they come from different sources:

-[Joe laughing hysterically]

-[Maria] I can't believe you did that!

-[Joe laughing hysterically]

-I can't believe you did that!

- If the sound effect emanates from the speaker themselves, no hyphens are needed:

[Joe laughing hysterically]

I can't believe you did that!

- Use hyphens to distinguish two distinct sound effects emanating from different sources:

-[horse neighs]

-[engine starts]

- Text in each line in a dual speaker subtitle must be a contained sentence and should not carry into the preceding or subsequent subtitle. Creating shorter sentences and timing appropriately helps to accommodate this.

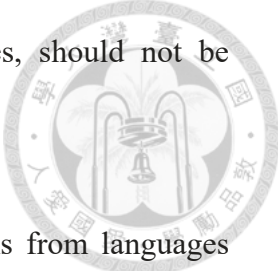
I.7. Font Information

- Font style: Arial as a generic placeholder for proportional SansSerif
- Font size: Relative to video resolution and ability to fit 42 characters across screen
- Font color: White



I.8. Foreign Dialogue

- In instances of foreign dialogue being spoken:
 - If foreign dialogue is translated, use [in language], for example [in Spanish]
 - If foreign dialogue is not meant to be understood, use [speaking language], for example [speaking Spanish]
 - Always research the language being spoken – [speaking foreign language] should never be used
 - Accents or dialects require the same treatment, for example [in Spanish accent]
- Foreign words that are used in a mostly English line of dialogue do not require identifiers, but should be italicized. Always verify spelling, accents and punctuation, if applicable.
- Familiar foreign words and phrases which are listed in *Webster's* dictionary should not be italicized and should be spelled as in *Webster's* dictionary (e.g. bon appétit, rendezvous, doppelgänger, zeitgeist, etc.).

- 
- Proper names, such as foreign locations or company names, should not be italicized.
 - Always use accents and diacritics in names and proper nouns from languages which use the Latin alphabet where their use is seen in official sources, or in the source text for fictional names. For example, Spanish names such as Mónica Naranjo, Pedro Almodóvar, Plácido Domingo should retain their diacritics. Any proper names which have lost the use of accents due to cultural reasons (e.g. Jennifer Lopez) do not need to have them added.
 - Transliterate uncommon or unfamiliar letters/characters which appear in names or proper nouns when working from a Roman alphabet language into English if they may cause confusion or be hard to understand or pronounce. Note that diacritics should be kept in proper nouns and names. For example: If the Icelandic name Þór appears, please transliterate as Thór (following relevant KNP and guidance about handling character names). If a German street name such as Torstraße appears in the source, please transliterate as Torstrasse (following relevant KNP and guidance about handling character names).

I.9. Italics

- Italicize text only in the following cases:

- 
- Narration
 - The voice of a visible character expressing unspoken thoughts or inner monologue
 - Song lyrics when sung, not quoted (if rights have been granted)
 - Unfamiliar foreign words and phrases which do not appear in the nominated dictionary for your language (do not italicize foreign loan words which appear in your language's nominated dictionary, e.g. rendezvous, zeitgeist etc. for English)
 - Proper names, such as locations, vessels names or company names, should not be italicized
 - Dialogue that is heard through electronic media, such as a phone, television (especially if we see the television and hear the audio), computer, loudspeaker, non-sentient robots, robotic voices or AI, etc.
 - In sections such as a phone conversation where the shot changes regularly between speakers, always ensure that segmentation and timing rules are correctly applied so as to ensure italics are used consistently and correctly
 - Only use italics when the speaker is not in the scene(s), not merely off screen, behind a door or out of shot

- 
- Titles of books, periodicals, works of art, albums, movies, TV shows, radio shows, video games, etc. (for an episode title in a series or song titles use quotation marks)
 - Only italicize titles, not names (e.g. the title of a book but not the name of a ship)
 - Italics may be used when a word is obviously emphasized in speech and when proper punctuation cannot convey that emphasis (e.g. *It was*).
 - In trailers, where dialogue rapidly switches between off-screen characters, on-screen characters and narrators, do not italicize any dialogue from the characters and speakers and only italicize narration.
 - This is the only set of rules to be followed for application of italics and trumps any additional advice found in associated references.

I.10. Line Treatment

- Maximum two lines.
- Text should usually be kept to one line, unless it exceeds the character limitation.
- Prefer a bottom-heavy pyramid shape for subtitles when multiple line break options present themselves, but avoid having just one or two words on the



top line.

- Follow these basic principles when the text has to be broken into 2 lines:
- The line should be broken:
 - after punctuation marks
 - before conjunctions
 - before prepositions
- The line break should not separate
 - a noun from an article
 - a noun from an adjective
 - a first name from a last name
 - a verb from a subject pronoun
 - a prepositional verb from its preposition
 - a verb from an auxiliary, reflexive pronoun or negation

I.11. Numbers

- From 1 to 10, numbers should be written out: *one, two, three*, etc.
- Above 10, numbers should be written numerically: *11, 12, 13*, etc.
- When a number begins a sentence, it should always be spelled out.
- Times of day:

- Use numerals when exact times are emphasized: *9:30 a.m.*
- Use lowercase *a.m.* (ante meridiem) and *p.m.* (post meridiem) when mentioned in dialogue
- Spell out words/phrases that do not include actual numbers: *half past, quarter of, midnight, noon*
- When *o'clock* is mentioned in dialogue, always spell out the number: *eleven o'clock in the morning*
- Note that the above rules may be broken due to space limitations or reading speed concerns, as well as for consistency when listing multiple quantities, for example.

I.12. Punctuation

- Avoid using complex punctuation which could be hard for viewers to follow. For example, avoid using colons and semi-colons and instead use simple, clear sentence structures to aid comprehension.
- Avoid over-use of commas (e.g. Prefer *I love you too* rather than *I love you, too*. Also *He told me too* rather than *He told me, too*).
- Double spaces are not permitted.
- En and em dashes are not permitted.



- Hash symbols may be used when someone mentions a hashtag. Spell out the word “hashtag” when used as a verb.
- Ampersands may be used when part of an initialism such as R&B or B&B.
- Use exclamation marks only in cases of shouting or surprise. Avoid over-using them.
- Interrobangs may be used in cases of a question being emphatically asked in an excited/shocked way or in disbelief. Prefer the format ?!, e.g. *What did you say?!*



I.13. Quotes

- Quoted words, phrases and sentences are enclosed in double quotation marks; single quotation marks enclose quotations within quotations.

He told me, "Come back tomorrow."

He said, "'Singing in the Rain' is my favorite song."

- If the quote extends beyond more than one subtitle, use an open quote at the beginning of the first subtitle, at the start and end of sentences within the quote and an end quote at the end of the last subtitle.

Subtitle 1 "Good night, good night!"

Subtitle 2 "Parting is such sweet sorrow

Subtitle 3 that I shall say good night till it be morrow."

- Use U.S. English rules:

- Periods and commas precede closing quotation marks, whether double or single.

- Colons and semicolons follow closing quotation marks.

- Question marks and exclamation points follow quotation marks unless they belong within the quoted text:

Which of Shakespeare's characters said, "Good night, good night"?

Juliet said, "Good night, good night!"

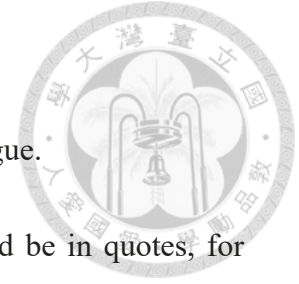
- Song titles should be enclosed in quotes.
- Use quotation marks when a character is seen to be reading aloud.
- If an on-screen character does “air quotes” when speaking, please apply quotation marks to the equivalent word in the target language in order to retain creative intent and to help ensure clarity about which word or part of the sentence the air quotes apply to.

I.14. Reading Speed Limits

- Adult programs: Up to 20 characters per second
- Children's programs: Up to 17 characters per second

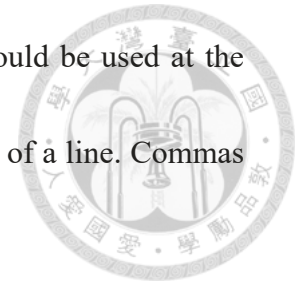


I.15. Songs



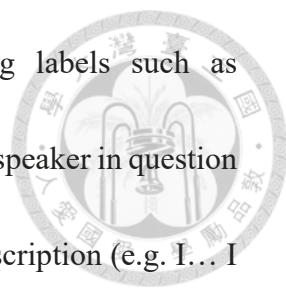
- Subtitle all audible song lyrics that do not interfere with dialogue.
- Use song title identifiers when applicable - song titles should be in quotes, for example ["Forever Your Girl" playing]. Use the name of a musical number or classical piece only if widely known, for example ["The Nutcracker Suite" plays].
- A song title can be added as an SDH label when there is room at the start of the song or when lyrics are not being included (e.g. if they remain untranslated as per the non-SDH experience or if dialogue and/or sound effects are taking precedence).
- Italicize lyrics.
- Song lyrics should be enclosed with a music note (♪) at the beginning and the end of each subtitle.
- Add a space between the music note and the preceding or subsequent text.
- When a dual speaker subtitle appears in a song, e.g. when there is a duet, each line of sung text should have a music note at the beginning and end to clearly indicate that both characters are singing.
- Use an uppercase letter at the beginning of each line including the first letter of the first word of the second line in a two-line subtitle
- Use ellipses when a song continues in the background but is no longer subtitled to give precedence to dialogue.

- Punctuation: only question marks and exclamation marks should be used at the end of a line - no commas or periods are to be used at the end of a line. Commas can be used within the lyric line, if necessary.
- Album titles should be in italics.
- Song titles should be in quotes.
- Follow this approach for poetry but do not include music notes.



I.16. Speaker ID / Sound Effects

- Use brackets [] to enclose speaker IDs or sound effects.
- Use all lowercase, except for proper nouns.
- Be detailed and descriptive, use adverbs where appropriate when describing sounds and music, describe voices, speed of speech, volume of sound.
- Describe the sounds and audio as opposed to visual elements or actions.
- Subtitle silence if plot-pertinent. For example, when plot-pertinent music ends abruptly.
- Plot-pertinent sound effects should always be included unless inferred by the visuals.
- Speaker IDs and the corresponding dialogue should ideally be on the same line.

- 
- When describing hesitations and nervousness, avoid using labels such as [stutters], [stuttering], [stammers] and [stammering] unless the speaker in question has a stutter/stammer. Instead, represent hesitations in the transcription (e.g. I... I said no!) or using sound labels such as [hesitates] or [spluttering], for example.
 - Only use speaker IDs or sound effects when they cannot be visually identified.
 - When a speaker ID is required for a character who has yet to be identified by name, use [man] or [woman], or [male voice] or [female voice], so as not to provide information that is not yet present in the narrative. If the same identifier is used multiple times in one scene, numbers should be added to distinguish them, for example [man 1].
 - Use a generic ID to indicate and describe ambient music, for example [rock music playing over stereo].
 - Use objective descriptions that describe genre or mood identifiers for atmospheric non-lyrical music, for example [menacing electronic music plays].
 - Sound effects should be plot-pertinent.
 - Sound effects that interrupt dialogue should be treated as follows:

Subtitle 1: However, lately, I've been...

[coughs, sniffs]

Subtitle 2: ...seeing a lot more of this.

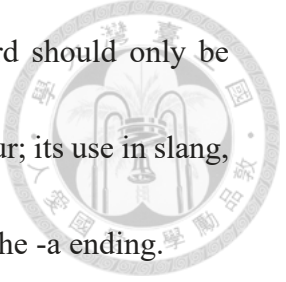
- Never italicize speaker IDs or sound effects, even when the spoken information is italicized, such as in a voice-over.

[narrator] *Once upon a time, there was...*



I.17. Special Instructions

- Dialogue must never be censored.
- Plot-pertinent dialogue always takes precedence over background dialogue.
- Deliberate misspellings and mispronunciations should not be reproduced in the translation unless plot-pertinent.
- When the word “black” appears in reference to someone’s race or ethnicity, capitalize it as Black. Use this form when referring to an African American or Black person, when referring to the African diaspora and when referring to collective groups or institutions, e.g. Black cinema, the Black community, a Black person. Please follow this rule when writing in or transcribing all variants of English. Always follow the word order and choice of the audio when working with SDH.
- Similarly, please capitalize the following words when used in reference to people and communities: Deaf, Indigenous

- 
- When transcribing usage within the English audio: the n-word should only be spelled with the -er ending in historical contexts or as a racist slur; its use in slang, non-racist conversation, or song lyrics should be handled with the -a ending.
 - Where variant spelling options present themselves, please opt for the most common spelling for the variant of English you are working with. Always prefer "okay", rather than "OK" or "Ok".

I.18. Reference

For all language-related issues not covered in this document, please refer to:

- The Merriam-Webster dictionary: <http://www.merriam-webster.com/>
- Chicago Manual of Style: <http://www.chicagomanualofstyle.org/home.html>

Change Log:

2022-12-22

- Revised section I.15 [Songs](#) - final bullet point expanded to confirm that music notes should not be included in poetry
- Revised sections I.14 [Reading Speed](#) - sections edited to mention "reading speed limits" and "up to"

2021-10-22

- Revised sections I.8 Foreign dialogue - bullet point added regarding transliteration of unfamiliar characters in proper nouns/names



2021-07-11

- Revised sections I.3 Continuity - 2nd bullet edited confirming that in the case of short or no pause, ellipses are not to be used
- Revised sections I.13 Quotations - 6th bullet point added
- Revised sections I.10 Line Treatment - 3rd bullet point added and line break information added
- Revised section I.15 Songs - 3rd, 6th and 7th bullet points added
- Revised sections I.12 Punctuation - 2nd and 4th bullet points added regarding dashes and comma over-use, 7th bullet added regarding exclamation marks, 8th bullet added regarding interrobangs
- Revised sections I.17 Special Instructions - final bullet about variant spellings added

2021-02-24

- Revised section I.16 Speaker IDs and sound effects - 3rd-7th bullet points added



2020-10-29

- Revised section I.1 [Accuracy of content](#) - 5th bullet point added
- Revised section I.4 [Dates and decades](#) - 4th bullet point added
- Revised section I.6 [Dual speaker subtitles](#) - 6th bullet point added
- Revised section I.10 [Line Treatment](#) - 2nd bullet point added
- Revised section I.13 [Quotations](#) - 5th bullet point added
- Revised section I.15 [Songs](#) - 10th bullet point added

2020-07-27

- Revised sections I.9 and [Italics](#) - whole section revised
- Revised section I.15 [Special Instructions](#) - 4th bullet reworded, 5th bullet added
- Revised sections I.3 and [Continuity](#) - 1st bullet added confirming type of ellipsis permitted, 3rd, 5th and 6th bullets revised
- Revised section I.1 [Accuracy of Content](#) - 3rd bullet reworded
- Revised section I.13 [Songs](#) - 5th and 7th bullets reworded
- Added sections I.4 [Dates and Decades](#) (subsequent sections renumbered)
- Added sections I.12 [Punctuation](#) (subsequent sections renumbered)

- Revised sections I.12 [Quotations](#) - 2nd bullet and corresponding examples updated



2020-06-16

- Revised section I.15 [Special Instructions](#) - 4th bullet added

2020-05-08

- Revised section I.1 [Accuracy of Content](#) - 3rd bullet about matching English dubs when creating SDH added
- Revised section I.7 [Foreign Dialogue](#) - 5th bullet point added about diacritics on non-English names
- Revised section I.14 [Speaker ID / Sound Effects](#) - 3rd bullet point added regarding representation of hesitation

2018-11-28

- Revised section I.1 [Accuracy of Content](#) - revised fourth bullet point
- Revised section I.7 [Foreign Dialogue](#) - 4th sub point added to 1st bullet point, 2nd and 3rd bullet points merged and reworded for clarity
- Revised section I.13 [Songs](#) - 2nd bullet point revised for clarity

- Revised section I.14 [Speaker ID / Sound Effects](#) - 4th bullet point revised for clarity and 6th bullet point added



2018-07-19

- Revised section I.1 [Accuracy of content](#) - 4th bullet added for dialectal treatment

2018-03-09

- Revised section I.3 [Continuity](#) - example added to 3rd bullet point
- Revised section I.4 [Documentary](#) - 1st bullet point revised, 2nd bullet point added
- Revised section I.5 [Dual Speakers/Multiple Events](#) - 2nd, 3rd, 4th and 5th bullet points added
- Section 7 [Forced Narratives](#) - moved to Section II
- Revised section I.7 [Foreign Dialogue](#) - 1st bullet point added
- Added section I.9 [Line Treatment](#)
- Revised section I.10 [Numbers](#) - 5th bullet point revised
- Revised section I.12 [Reading Speed](#) - words per minute removed
- Revised section I.14 [Speaker ID / Sound Effects](#) - 4th and 7th bullet points added

2016-05-15

- Revised section 10 Numbers - 4th bullet point revised for clarity
- Revised section 13 Songs - 8th bullet point revised
- Revised section 14 Speaker ID / Sound Effects - 7th bullet point, 3rd item added



2015-12-07

- Revised section 7.13 Songs - 1st bullet point revised, 2nd bullet point added