Towards a corpus-based approach to graphic elements in creative subtitling:

A case study of a YouTube channel "Apenjie with Dawang"

Abstract:

This article attempts to explore how a corpus-based approach allows us to describe and analyze the multimodal complexity of graphic elements in creative subtitling. To this end, the article focuses on a YouTube channel, *Apenjie with Dawang*, featuring a dog and his owner. In this channel, the subtitling strategies were experimental with multiple graphic elements (colors, positions, font sizes, and emojis). Informed by a social semiotic approach to multimodality, a corpus of 1,155 coupled pairs of Chinese-English subtitles were annotated for modal shifts and metafunctional shifts. Some major findings include: (a) emojis were much more likely to be added to the target subtitles for the animals than for the humans; (b) speaker-identifying graphic elements (color and position) were lost in the target subtitles, but emojis were systematically added to mark animals as the speakers; (c) the addition of emojis evoked complementary-interpersonal meanings, suggesting that the subtitlers might have prioritized audience engagement over textual fidelity; (d) although the target subtitles used fewer graphic elements, the semiotic meanings could be similar or complementary to those of the source subtitles. Based on these findings, the article also discusses the opportunities and challenges of a corpus-based approach to graphic elements in creative subtitling.

Keywords: creative subtitling; corpus; graphic elements; multimodality; metafunction; mode

1. Introduction

Multimodality has garnered sustained attention from researchers on audiovisual translation (AVT) (Kruger 2021; Remael and Reviers 2020; Taylor 2013, 2016, 2020). Although multimodality has been a major concern in AVT scholarship, previous studies have tended to focus on practices on traditional platforms, such as subtitling and dubbing for films and TV series. With respect to subtitling practices, scholarly attention has been primarily given to conventional subtitles, typified by a singular language mode, a fixed placement, a uniform color, and an invariable font size (Díaz Cintas and Remael 2021). By contrast, a small but growing body of research has begun to explore creative subtitles[1] that are unbounded by subtitling conventions about positions (Fox 2016, 2018), colors (Foerster 2010; McClarty 2012, 2014), spellings (Secară 2011), and other graphic features (Díaz Cintas 2018). This article attempts to extend the research focus to include one more graphic element in creative subtitling: emojis. As illustrated in Figure 1, the source Chinese subtitle is different from standard subtitles in terms of the position (close to the "speaker" vs. fixed at the bottom), color (yellow vs. white), and font design (varying sizes vs. an invariable size). The target English subtitle line (albeit rendered in a standard position, color, and font size) is adorned with an emoji to mark the playful tone. These dynamic shifts of graphic elements are increasingly popular on streaming platforms, where subtitlers harness a wider range of multimodal resources to attract and engage viewers (Díaz Cintas 2018). As such, creative subtitling practices raise the questions of whether semiotic meanings have been altered due to the shifts of graphic elements, and if so, how. Given the fluidity of creative subtitling practices on streaming platforms, multimodality is better approached from a social semiotic perspective that "deals with *meaning* in

---

[1] While there are other terms to describe unconventional subtitling practices, such as "abusive subtitles" (Nornes 1999) and "free form subtitles" (Bassnett et al. 2022), this article follows McClarty (2012) and uses the term "creative subtitles" to highlight the creative deployment of graphic elements in the subtitling practice "that differs from the norm" (McClarty 2012: 139).

all its appearances, in all social occasions and in all cultural sites" (Kress 2010: 2, original emphasis). In addition, previous studies have tended to focus on a limited number of cases of creative subtitling (see the review below), with relatively little attention to the patterns emergent from multiple audiovisual texts. In this regard, a corpus-based approach allows us to systematically examine the patterned shifts of graphic elements and their effects on semiotic meanings. The corpus-based observations may afford interesting insights that are beyond the remit of one case or several cases. In the following sections, I will first explain a semiotic approach to multimodality as a theoretical framework and a corpus-based approach as a research methodology. With these conceptual and methodological tools, I will present a corpus-based case study on a YouTube channel featuring the creative use of graphic elements. Building on the research findings, I will discuss the opportunities and challenges of a corpus-based approach to graphic elements in creative subtitling.

Figure 1. Example of creative subtitles (produced by the author)

2. Research background

2.1 A social semiotic approach to multimodality

A social semiotic approach to multimodality draws on Hallidayan Systematic Functional Grammar (SFG) and highlights that meaning making is contingently shaped by the social context and the designer's agency (Jewitt, Bezemer and O'Halloran 2016). Semiotic resources are selected and orchestrated in ways that the designer considers apt for the communicative context (Kress 2010). Fundamentally, semiotic resources are organized to signify meanings and fulfil three metafunctions: *ideational*, *interpersonal*, and *textual* (Halliday 1978; Halliday and Matthiessen 2014; Kress and van Leeuwen 2006). The *ideational* metafunction refers to the representation of "human experience" (Halliday and Matthiessen 2014: 30), such as participants (human or non-human agents), actions, and states. The *interpersonal* metafunction alludes to relations enacted by semiotic resources, typically involving three types of relations: (a) "between represented participants"; (b) "between interactive and represented participants"; and (c) "between interactive participants" (Kress and van Leeuwen 2006: 114). In the case of AVT, represented participants can be a person and a dog in a film, while interactive participants can be the audience and the producer of the film. At the service of the two previous metafunctions, the *textual* metafunction means that semiotic resources are coherently organized to enable the creation of ideational and interpersonal meanings. Applied to Translation Studies, the metafunctional analysis is heuristically useful in examining whether and how ideational, interpersonal and textual meanings are shifted between the source and the target texts (see Halliday 2001; Matthiessen 2021; Tuominen, Jiménez Hurtado and Ketola 2018). There have been some attempts to conduct the metafunctional analysis of AVT products, such as subtitling (Chen 2019; Mubenga 2009), audio description (Reviers 2018), and dubbing (Pérez-González 2007). Some earlier studies have demonstrated the descriptive and

explanatory power of the metafunctional analysis using a case-based design (Mubenga 2009; Pérez-González 2007; Taylor 2003). However, it is not until recently that AVT researchers have combined a corpus-based design with a systematic-functional lens to reveal patterns across a larger number of samples (Chen 2019; Reviers 2018).

2.2 A corpus-based approach to subtitles

Over the past decades, an increasing number of studies have opted for a corpus-based[2] design to examine AVT products, such as subtitling (e.g. Tirkkonen-Condit and Mäkisalo 2007; Xavier 2022), dubbing (e.g. Romero-Fresco 2009; Valentini 2007) and audio description (e.g. Perego 2019; Reviers 2018). Given the research focus in this article, the following paragraphs will review the corpus-based studies on subtitling only. Interested readers may refer to reviews of corpus-based studies on other AVT modalities elsewhere (e.g. Bruti 2020; Pavesi 2019).

Existing corpus-based studies on subtitling can be broadly grouped into two strands, corresponding to two types of corpora built by the researchers: comparable corpora and parallel corpora[3]. The first strand examines the linguistic and stylistic features of subtitles by comparing the translated texts and their domestic counterparts. For instance, Tirkkonen-Condit and Mäkisalo (2007) contrasted the cohesive devices in Finnish subtitles with those in Finnish translations and native Finnish writings. They found that the frequency of cohesive devices differed among the three corpora, suggesting that the subtitled language might be a distinct language variety. More recently, Levshina (2017) compared the original English subtitled language, the translated English

---

[2] While a distinction has been made between *corpus-based* and *corpus-driven* approaches (e.g. Tognini-Bonelli 2001), some scholars challenge this distinction (e.g. McEnery and Hardie 2012) and believe that "all corpus linguistics studies are necessarily corpus-based" (Brookes and McEnery 2020: 382). The debate is beyond the scope of the current study. Throughout this article, *corpus-based* is used in its broad sense, denoting studies that draw on corpus data in varying degrees (McEnery and Hardie 2012).
[3] Comparable corpora are "paired on the basis of textual similarity" (Zanettin 2012: 10) and may be monolingual or multilingual, while parallel corpora are those with source texts and corresponding target texts (see also Pavesi 2022).

subtitled language, and the native English spoken and written registers. Based on the n-gram distribution, the analysis revealed that the subtitled languages (original or translated) approximated the features of native English informal conversations. The analysis also showed that the translated subtitles contained less discourse markers and tended to be more formal and more narrative than the original English subtitles.

Different from the first strand, the second strand of research draws on parallel corpora to describe subtitling strategies and shifts between source and target texts. As an example, Bywood (2019) built a corpus of multiple English-subtitled versions of two German films. The analysis focused on the subtitling strategies of culture-specific references in the English versions that spanned about a decade. The results showed that the percentages of source-oriented strategies (e.g. direct translation) decreased, while the percentages of target-oriented strategies (e.g. substitution) increased as the time went by. Similarly, Xavier (2022) drew on a corpus of six English movies and their Portuguese subtitles to investigate how taboo language was rendered. It was found that over three quarters of taboo words were neutralized (i.e. omitted, standardized or euphemized) in the target subtitles.

The corpus-based studies reviewed previously have tended to focus on the subtitled texts, with relatively less attention to multimodality. While these linguo-centric studies have shed important light on the recurrent features of the subtitled language and the patterns of translators' decisions, they do not fully account for the meanings afforded and constrained by the multiplicity of semiotic resources. Different from the primary focus on the language mode, a small but growing body of research has taken a corpus-based approach to examining subtitles as part of the multimodal whole. As a recent example, Pinto and Mubaraki (2020) have shown how the standardization of sociolects in subtitling may change the intermodal relations among the subtitle, speech, and mise-en-scène.

6

The corpus-based analysis also revealed that an altered intermodal relation may or may not lead to the change of diegetic functions, depending on whether the intermodal relation in the source text was of contradiction or confirmation. As such, a corpus-based approach to multimodality in subtitling enables us to "deal with the complex semiotic fabric of audiovisual texts" (Tuominen, Jiménez Hurtado and Ketola 2018: 6). This approach is particularly relevant to the current research on graphic elements in creative subtitling, with a notable advantage of "reveal[ing] cumulative effects that would not emerge through the examination of individual translations" (Pavesi 2019: 315). In other words, unlike previous case-based studies on creative subtitling (e.g. Foerster 2010; McClarty 2012), the current research will examine the frequency of multimodal features across multiple audiovisual texts so that general tendencies will become visible. In doing so, we will be able to understand the patterns inherent in texts, translators, and multimodal phenomena that are difficult to be captured by the intuition of researchers (Baños, Bruti and Zanotti 2013; Soffritti 2019). Therefore, the current study adopts a corpus-based approach to multimodality, informed by the social semiotic approach as the theoretical framework. A multimodal corpus was built to examine graphic elements in creative subtitling and map the general trends of meaning shifts in relation to modal transformations.

3. The study

3.1 Corpus

This study focused on a YouTube channel, *Apenjie with Dawang* (a derivative of the popular channel *Dianxi Xiaoge*). The channel was chosen because the subtitling practices were experimental and exemplary of the creative subtitling strategies increasingly popular on digital platforms (Díaz Cintas 2018). The channel features a dog *Dawang* and his owner *Apenjie*. Chinese

intralingual subtitles (adorned with the graphic elements of colors, positions and font sizes) are integrated into the visuals to (a) give a voice to the dog for comical effects; and (b) provide Mandarin-Chinese subtitles for viewers to understand the dialogues in a local dialect. In contrast, the English subtitles are rendered as a combination of verbal texts and emojis. Interestingly, the percentage of the emojis in the English subtitles is much higher than that in the Chinese subtitles. In addition, the numbers of modes deployed in the source and target subtitles differed. These creative semiotic features present rich opportunities for us to understand the patterns and diversity of graphic elements in creative subtitles.

By mid-November 2021, the time when the data were collected, the channel released 127 videos. To determine the number of videos included for the corpus, a data saturation method was used (see Miles and Huberman 1994; Hart 2020). Specifically, at the outset, a random sample of 10 videos were selected, annotated, and analyzed for patterns (detailed in Section 3.2). Then, another random sample of 10 videos were added and the processes were repeated until "adding to the sample no longer provide[d] any substantive new insight" (Hart 2020: 152). As it turned out, the inclusion of a total of 30 videos stabilized the patterns observed in the corpus (see Section 3.3 for detail). Thus, the resultant corpus consisted of the source and target texts from these 30 videos.

3.2 Coding framework

The corpus was annotated with a coding framework informed by the social semiotic approach to multimodality. The framework covered two broad dimensions: subtitling and sound, as shown in Figure 2. The braces indicate non-exclusive coding options, such as the possible co-presence of language, color and font size as modal choices. The square brackets mean that the codes are exclusive and only one code can be selected for a particular case. The two-sided arrows mean that

the actual codes are possible combinations of the two sub-categories, such as "complementary-interpersonal" and "divergent-ideational" (described further below).
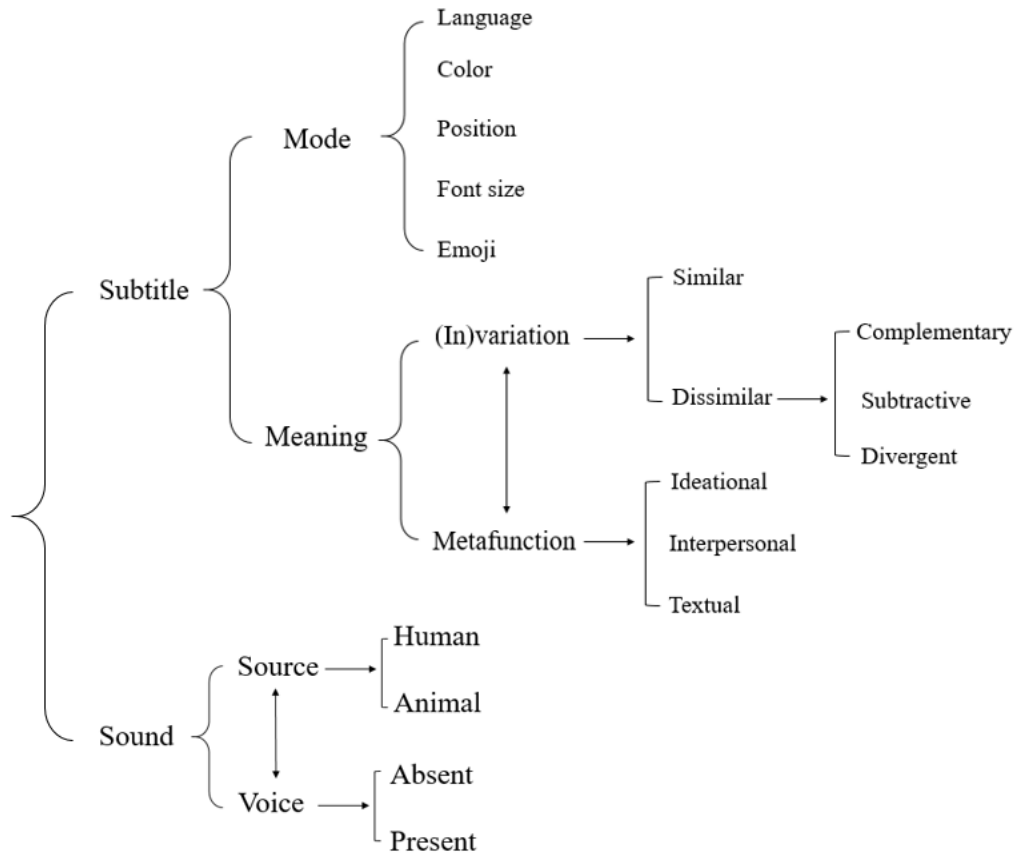

Figure 2. Coding framework

The first dimension about creative subtitling focused on two main aspects: modes and meanings. Specifically, the framework included five modes as meaning-making resources in the source and target texts: language, color, position, font size, and emoji. These five modes were included for annotation because they represented the full range of graphic elements for the subtitling designs on the focal YouTube channel. While it is straightforward to regard language and emoji as two distinctive modes, the modal status of color, position, and font size warrants

justification. The social semiotic approach to multimodality maintains that "a mode is what a community takes to be a mode and demonstrates that in its practices" (Kress 2014: 65). On digital platforms and among fansubbing communities in particular, the color, position, and font size have a wide range of meaning potentials, such as identifying speakers, enhancing characterization, and marking emotions (Foerster 2010; McClarty, 2014; Pinnow 2015). Therefore, these semiotic resources can be considered as modes to fulfill representational needs (see Kress 2014).

In addition to modal composition, the framework examined whether and how the semiotic meanings shifted between the source and target subtitles (represented by the sub-categorical codes under "Meaning" in Figure 2). The three metafunctions reviewed in Section 2.1 guided the analysis. The coding labels were inspired by Kress (2020) to describe meaning (in)variations. If the target subtitles simultaneously fulfilled similar ideational, interpersonal, *and* textual metafunctions as those in the source subtitles, they had *similar* semiotic meanings. Regarding *dissimilar* scenarios, three were differentiated: (a) if the target subtitles carried more contextually-congruent meanings than the source subtitles in any of the metafunctions, the translated texts had *complementary* semiotic meanings; (b) if the target subtitles carried fewer meanings than the source subtitles in any of the metafunctions but the meanings were still contextually-congruent, the translated texts had *subtractive* semiotic meanings; (c) if the target subtitles carried contextually-incongruent meaning to the source subtitles in any of the metafunctions, then the translated texts had *divergent* semiotic meanings. In scenarios (a), (b) and (c), the affected metafunctions were also specifically coded for subsequent analysis. Based on three rounds of coding, five codes emerged: *similar*, *complementary-interpersonal*, *subtractive-interpersonal*, *divergent-ideational*, and *divergent-textual* (more on this in Section 3.3).

The second dimension of the framework aimed to capture whether and how subtitling choices were influenced by the sound, since the focal YouTube channel featured frequent dialogic interaction between the dog and his owner. Two aspects of the sound were coded: source and voice. Specifically, subtitles might or might not co-occur with the human voices (e.g. talking and shouting) or the animal voices (e.g. barking and woofing). These were represented by the combination of the sub-categorical codes "source" and "voice" as follows (see also Figure 2): *human-voice-present*, *animal-voice-present*, *human-voice-absent*, and *animal-voice-absent*.

The following examples illustrate the coding framework in greater detail (see Figure 3). These examples are reproduced from the videos and the corresponding video links can be found in the Reference. In the first example, the source subtitles co-occur with the human voice. Three modes (language, color, and position) are used to signify the speaker and the speech content. By contrast, the target subtitles solely rely on the language mode to represent the speech content, without any modal resource to identify the speaker. However, since only one person is talking at the time, the speaker's identity is clear from the audio input. Thus, the target subtitles do not differ from the source ones in the three metafunctions, and are considered *similar* in terms of semiotic meanings.

Examples*	Sound	Source subtitles	English gloss of the source subtitles	Source-subtitle modes	Target subtitles	Target-subtitle modes	Semiotic meanings
1	Human-voice-present	你看什么 你又不会绑	What are you looking at; you don't know how to tie	Language, color, position	what are you looking at, you don't know how to wrap this~	Language	Similar
2	Animal-voice-absent	可怜	Pathetic	Language, color, position	🥺🥺🥺	Emoji	Divergent-ideational
3	Human-voice-absent	紧张 😫	Nervous	Language, color, position, emoji	nervous…	Language	Subtractive-interpersonal
	Animal-voice-present	给我点恰恰呀~	Give me some to eat ah~	Language, color, position,	(give me some food~ 😳)	Language, emoji	Complementary-interpersonal
4	Animal-voice-absent	嘿嘿	Hei Hei**	Language, color, position	😁😁😁	Emoji	Complementary-interpersonal
	Human-voice-present	游一下（小朋友的声音）	Swim a bit (children's voice)	Language, color, position	(swim a little bit~) from the kids	Language	Divergent-textual

* The corresponding video links of these examples can be found in the Appendix.

** In Chinese, *Hei Hei* is an onomatopoeia of the laughing sound.

Figure 3. Examples of the corpus annotation

In the second example, the source subtitles are for the dog (Dawang) unaccompanied by his voice. Three modes (language, color, and position) are used to signify Dawang's speaking turn and his speech content. The target subtitles use emojis only. In the previous scenes, some food was given to a kid, while Dawang got minimal "left-over" food, so the source subtitles characterized him as "pathetic", denoting a sad, pitiful state. The target subtitles, however, use three "Face with Steam from Nose"[4] emojis. The emoji, when used for negative emotions, usually signifies "irritation, anger, and contempt" (https://emojipedia.org/face-with-steam-from-nose/). As these feelings are incongruent with the emotional state (sadness) represented in the source subtitles, the instance is coded to have a *divergent-ideational* meaning.[5]

Example 3 consists of two parts of subtitles simultaneously presented on the screen. As the first part, the source subtitles signify the speaking-turn of the grandmother, but she does not make any sound (human-voice-absent). The subtitles are rendered in a white color and positioned above her head. The subtitle text reads *jinzhang* (nervous) followed by a "Downcast Face with Sweat" emoji that intensifies the mood. However, the emoji is lost in the target text that now relies on the language mode only. As emojis are "mood enhancers" (Danesi 2017: 97), it can be argued that the removal of the emoji makes the mood less pronounced to the audience. Thus, the instance is coded to have a *subtractive-interpersonal* meaning. The second part of the source subtitles are for the dog, accompanied by his burping. The yellow-color subtitles are positioned above his head. Different from the source subtitles, the target subtitles do not use color or position for speaker identification.

---

[4] In this article, short codes from Emojipedia (https://emojipedia.org/) are used to describe the emojis.

[5] In SFG, this is tantamount to a meaning shift of the Attribute in the Relational Process (Halliday and Matthiessen 2014).

Instead, the focal YouTube channel resorts to two typical multimodal strategies. First, emojis are systematically added to the target subtitles to mark the animal's "speaking" turns (see Examples 2, 3 and 4 in Figure 3 as quick illustrations). Second, when dual speakers are presented within the same line of the target subtitles, the animal-speaker identity is further marked by the brackets. As in Example 3, brackets and a "Pleading face" emoji are added to the target subtitles for the dog to indicate his speaking turn. Relatedly, the addition of the emoji concretizes the emotion of the dog, thus evoking a *complementary-interpersonal* meaning (see also Logi and Zappavigna 2021 for the dual functions of emojis as discourse and emotive markers).

Example 4 is also comprised of two parts of subtitles simultaneously shown on the screen. The first part is the subtitles for the dog, *hei hei* (an onomatopoeia of the laughing sound), but actually he does not make any sound. The subtitles are yellow-colored, and positioned close to his head, thus making use of the modes of color, position and language. The target subtitles use the emoji mode only, more specifically a string of three emojis of "Beaming Face with Smiling eyes". Given the "tone-enhancing function" of the emojis (Danesi 2017: 96), this instance is coded to have a *complementary-interpersonal* meaning. The second part is the children's voice, represented by white-color subtitles at the bottom of the screen. Since the children are off-screen, round brackets are used in the source subtitles to explain who are making the remark "swim a bit". However, in the target subtitles, "swim a bit" is bracketed, while "from the kids" is un-bracketed. These are confusing because the audience may think that "from the kids" is the remark itself. Since the speaker identity and the speech content are not coherently represented, this instance is coded to have a *divergent-textual* meaning.

3.3. Coding and analytical procedures

Before the manual annotation, the source and target subtitles were aligned as "coupled pairs" (Bywood 2019; Toury 2012) according to the speaking turns (the units of analysis). The parallel subtitles were also time-stamped insofar as they could be located at particular moments of the videos, thus providing immediate access to the visual and aural channels (Thompson 2022). The coding process consisted of three steps. First, each video was watched in its entirety to get a sense of the multimodal whole. After this, each original-translated subtitle couplet was examined with its accompanying visual and/or aural information. Then, the coupled pair was coded based on the annotation framework outlined previously. When in doubt, the whole video was re-watched to inform the annotation.

As explained in Section 3.1, a data saturation method was used to determine the number of videos to be coded and analyzed (see Miles and Huberman 1994; Hart 2020). After the first 10 random videos were annotated, the distribution percentages of semiotic meanings were counted, with the total number of coupled pairs as the denominator (see Table 1). Then, the same coding and analytical procedures were repeated for the second 10 random videos. When the ranks of percentages did not change by adding new samples, the process stopped. As shown in Table 1, after the third round of coding (30 videos), the ranks of the semiotic meanings remained invariable—*complementary-interpersonal* and *similar* meanings consistently emerged as the top two categories, while *divergent-ideational*, *subtractive-interpersonal*, and *divergent-textual* meanings consistently ranked among the bottom three. Therefore, the patterns were considered saturated and the coding process stopped. To ensure coding reliability, a second coder annotated 268 coupled pairs from six

videos, respectively representing 23.2% and 20% of the total corpus samples. The kappa coefficient of 0.957 indicated a very strong inter-coder reliability.

Table 1. Cumulative percentages of semiotic meanings in the target subtitles

Meanings	10 videos (406 coupled pairs)	20 videos (786 coupled pairs)	30 videos (1,155 coupled pairs)
Complementary-interpersonal	272 (67%)	491 (62.5%)	694 (60.1%)
Similar	116 (28.6%)	243 (30.9%)	386 (33.4%)
Divergent-ideational	15 (3.7%)	34 (4.3%)	48 (4.2%)
Subtractive-interpersonal	3 (0.7%)	15 (1.9%)	20 (1.7%)
Divergent-textual	0	3 (0.4%)	7 (0.6%)

4. Results

The videos consisted of two groups of subtitles: human and animal. A total of 322 coupled pairs of subtitles were for the humans and 833 were for the animals. This distribution, albeit disproportional, was natural because the YouTube channel primarily featured Dawang, the dog. In the following sections, I will first describe the modal choices and semiotic meanings within each group, before making comparisons to examine whether inter-group differences exist.

4.1 Within-group comparisons

As shown in Figure 4[6], the target subtitles for the humans had a limited range of modal choices, largely relying on the language mode. In terms of modal shifts, the predominant strategy was to drop two modes (color and position) but preserve the language mode (298 instances out of 322). In the source subtitles, colors and positions were used for speaker identification. Despite the loss of these modes in the target subtitles, the semiotic meanings were chiefly similar to those of the source subtitles (294 instances), possibly because the vast majority of the subtitles co-occurred with the human voices (286 out of 294 instances). The aural input could signal to the viewers who was talking, thus serving similar semiotic functions of the colors and positions in the source subtitles (discussed further in Section 5.1 below about multimodal redundancy).

Modal shifts from the source to the target subtitles					Meaning (in)variations between the source and the target subtitles					Subtotal
Color	Position	Emoji	Font size	Language	Similar	Complementary-interpersonal	Subtractive-interpersonal	Divergent-ideational	Divergent-textual	
-	-			=	294 (286)*				4 (3)	298 (289)
-	-	-		=			15 (13)			15 (13)
-	-	+		-		2 (1)		1 (1)	1 (1)	4 (3)
-	-	+		=		3 (3)				3 (3)
-	-		-	=			2 (2)			2 (2)
				Subtotal	294 (286)	5 (4)	17 (15)	1 (1)	5 (4)	322 (310)

-	Removed
=	Preserved
+	Added

* Un-bracketed number: total frequency of the target subtitles *accompanied and unaccompanied* by the human voices

(Bracketed number): frequency of the target subtitles *accompanied* by the human voices

Figure 4. Descriptive statistics of the target subtitles for the humans

[6] In Figure 4 and Figure 5, the un-bracketed number within each cell represents the total frequency of the target subtitles *accompanied and unaccompanied* by the human/animal voices. The bracketed number represents the frequency of the target subtitles *accompanied* by the human/animal voices. For instance, the cell "294 (286)" in Figure 4 means that (a) 294 instances of the target subtitles were observed for the modal shifts (dropping the color and the position but preserving the language mode) conveying similar semiotic meanings; (b) these 294 instances comprised the ones co-occurring with the human voices (286 instances) and those unaccompanied by the human voices.

Regarding the subtitles for the animals (Figure 5), the target texts mainly employed the language mode and the emoji mode, while the source texts frequently utilized the modes of position, color and language. The analysis of modal shifts showed that the predominant strategy was to drop two modes (color and position) but preserve the language mode and add the emoji mode (575 instances out of 833). The second most commonly used strategy for the modal shifts (although a distant second, with 96 instances) was to remove three modes (color, position and language) but add the emoji mode. Specifically, the language texts in the source subtitles were replaced by emojis in the target subtitles. Most of the times, the emojis matched the ideational meanings of the substituted texts, such as the word *ku* (cool) being replaced by the "Smiling Face with Sunglasses" emoji. Despite the reduced number of modal resources, interpersonal meanings were in fact enhanced because emojis were stronger markers of emotions and intentions, as compared with their non-graphic textual counterparts.

Emojis were systematically added to the target subtitles for at least two functions. First, the subtitlers might attempt to further anthropomorphize the animals by associating them with emojis, most of which were human-faced and more explicit in the intended emotions. As such, the semiotic meanings of the target subtitles were primarily complementary (689 instances), while the similar, subtractive and divergent meanings were much less common. A second function of the emojis was to mark speaker identity. As shown in Figure 5, the target subtitles were rarely accompanied by the animal voices (34 out of 833 instances in total). With the removal of colors and positions for speaker identification and the absence of aural input, emojis were used as a marker to signal that the target subtitles were for the

animals. This could explain why emojis were rarely used in the target subtitles for the humans but lavishly added to those for the animals.

Modal shifts from the source to the target subtitles					Meaning (in)variations between the source and the target subtitles					Subtotal
Color	Position	Emoji	Font size	Language	Similar	Complementary-interpersonal	Subtractive-interpersonal	Divergent-ideational	Divergent-textual	
-	-	+		=		552 (19)*		23		575 (19)
-	-	+		-		84 (4)		11	1	96 (4)
	-	+		=		49 (4)		3		52 (4)
-	-	=		=	41 (2)			5		46 (2)
-	-			=	26 (3)			2	1	29 (3)
-	-	=		-	14			2 (1)		16 (1)
-	-	+	-	-	5 (1)					5 (1)
	-	+		-		4		1		5
-	-	+	-	=	4					4
-	-	-		=			2			2
-	-		-	=			1			1
	-	=		-	1					1
	-			=	1					1
				Subtotal	92 (6)	689 (27)	3	47 (1)	2	833 (34)

-	Removed
=	Preserved
+	Added

* Un-bracketed number: total frequency of the target subtitles *accompanied and unaccompanied* by the animal voices

(Bracketed number): frequency of the target subtitles *accompanied* by the animal voices

Figure 5. Descriptive statistics of the target subtitles for the animals

4.2 Inter-group comparisons

To examine whether the modal shifts and semiotic meanings were associated with the human or animal group of subtitles, Chi-square tests or Fisher's exact tests were performed where appropriate[7], followed by the calculation of odds ratios for focused comparisons.

With respect to the modal shifts, all colors, positions and font sizes were lost in the target subtitles, regardless whether they were for the humans or the animals. Therefore, the analysis focused on the language mode and the emoji mode, as they varied between the two groups. Table 2 reports the frequency of the language mode being removed from or

---

[7] If more than one fifth of expected counts fall below five, the Fisher's exact test is conducted, instead of the Chi-square test (see Field 2018).

preserved in the target subtitles. A significant association was found between the groups of target subtitles and the treatment of the language mode: $\chi^2(1) = 43.40$, $p < 0.001$. The odds ratio showed that the odds of having the language mode removed from the target subtitles were 13.77 times higher for the animals than for the humans.[8]

Table 2. Treatment of the language mode in the target subtitles (frequency count)

Group	Language mode removed	Language mode preserved
Animal	123	710
Human	4	318

Table 3 reports the frequency of the emoji mode being removed from, preserved in, or added to the target subtitles. The Fisher's exact test showed that a significant association was found between the groups of target subtitles and the treatment of the emoji mode ($p < 0.001$). Comparing the modal choices of having emojis added or removed, the odds ratio revealed that the odds of having emojis added to the target subtitles were 789.64 times higher for the animals than for the humans.[9]

Table 3. Treatment of the emoji mode in the target subtitles (frequency count)

Group	Emoji removed	Emoji preserved	Emoji added
Animal	2	63	737
Human	15	0	7

---

[8] The odds ratio was calculated as follows: (123/710) / (4/318) = 13.77.
[9] Given a 2×3 table and one cell in the second row with zero count, the odds ratio was calculated based on two columns (emoji added and emoji removed): (737/2) / (7/15) = 789.64 (see also Agresti 2013 for detailed explanation of calculating odds ratios for $I \times J$ tables).

Table 4 summarizes the frequency of (non-)shifts of semiotic meanings in the target subtitles. A significant association was found between the groups and the (in)variations of semiotic meanings: $\chi^2(4) = 757.15$, $p < 0.001$. The odds ratios in Table 4 were calculated with the *similar* column as the benchmark. As can be seen, the odds of enacting the *complementary-interpersonal* and *divergent-ideational* meanings in the target subtitles were much higher for the animals than for the humans. This points to two patterns: (a) the target subtitles primarily foregrounded the animal emotions and backgrounded the human emotions; (b) to enhance the animal emotions, a larger number of emojis were added, but the connoted emotional states might be inconsistent with the context (see Example 2 in Figure 3), resulting in higher odds of *divergent-ideational* meanings. For the two remaining shifts of semiotic meanings (*subtractive-interpersonal* and *divergent-textual*), the odds ratios were modest, possibly because the frequency counts were small in both groups.

Table 4. Frequency of the (in)variation of semiotic meanings in the target subtitles

Group	Similar	Complementary-interpersonal	Subtractive-interpersonal	Divergent-ideational	Divergent-textual
Animal	92	689	3	47	2
Human	294	5	17	1	5
	Odds ratios	440.36	0.56	150.20	1.28

5. Discussion

The corpus findings have generated some interesting insights into the graphic elements in creative subtitling. These will be discussed in Section 5.1, based on which Section 5.2 will discuss the opportunities and challenges of a corpus-based approach to creative subtitling.

5.1 Graphic elements in creative subtitling

As observed in this study, creative subtitling was liberal with the use of graphic elements (i.e. colors, positions, font sizes, and emojis), which arguably transformed the act of *reading* subtitles as monomodal texts into *viewing* subtitles as multimodal designs. In fact, the source and the target subtitles drew on different modal resources. Colors, positions and font sizes were primarily used in the source subtitles, while emojis were generously used in the target subtitles. These shifts of graphic elements lend further support to the "reductive" nature of subtitling (Dwyer 2017). In conventional subtitling, to cope with the time and space constraints, subtitlers may use the strategies of omission or condensation leading to loss in target subtitles (Pedersen 2011), such as the removal of verbal cues and sociolect markers. Different from these reductions of linguistic features, creative subtitling in this study was subjected to modal reductions. However, it should be noted that it was reductive in the number of modes, but not necessarily in the metafunctional meanings. As shown in Table 1, a substantial number of original-translated coupled pairs conveyed similar metafunctional meanings and an even larger number of target subtitles evoked complementary-interpersonal meanings.

Similar semiotic meanings were realized possibly due to multimodal redundancy (see also Remael and Reviers 2019). The varying positions of the source subtitles enacted

a *projection* function (i.e. assigning words to visual entities), typically realized through speech/thought bubbles in comics (Bateman 2014; see also Beckman 2008, who compared the creative subtitles in the film *Slumdog Millionaire* to dialogues in comics). The positioning of source subtitles worked like a speech bubble in comics, aligned with the proximity principle (Cohn 2013).[10] The adjacency of the subtitles and the visual entities (humans or animals) enabled the viewers to identify the speaker. As such, the three modes (language, position and image) were co-contextualized to signify who said what. However, in the target subtitles, the position mode and the concomitant projection function (speaker identification) were lost. In fact, another speaker-identifying graphic element (color) was also lost in the target subtitles. Despite the absence of these modes, they were largely redundant with the information from the soundtrack, image, and/or subtitle text. Thus, to a greater or lesser extent, viewers could infer the identity with the aids of (a) the speaker's voice that co-occurred with the subtitles; (b) the talking face in the visual frame; and (c) the dialogue content itself, such as pronouns and addresses (see Dahne and Piazza 2020; De Linde and Kay 1999). However, to further complicate the multimodal intricacies, the target subtitles for the animals were rarely accompanied by the animal voices (i.e. minimal aural input for the viewers to infer who was "talking"). As a possible solution to this issue, emojis were systematically added to the target subtitles as the markers of the animals to better facilitate speaker identification/differentiation. This could explain the significantly higher odds of emojis used in the target subtitles for the animals than for the humans.

---

[10] In Cohn's (2013) framework, a speech/thought bubble has three elements: *carrier* (place-holder of the text), *tail* (pointing to the visual entity), and *root* (the visual entity associated with the text). Creative subtitles with varying positions are tantamount to speech/thought bubbles without the carrier and tail, but with the text and the root (see Cohn 2013 for details).

Another possible reason for the lavish use of emojis might be related to the subtitlers' prioritized agenda to strengthen the rapport with the viewers. Given the "tone-enhancing function" of emojis (Danesi 2017: 96), the copious addition of emojis was more apt to evoke complementary-interpersonal meanings in the target subtitles for the animals. These semiotic shifts can be interpreted as the subtitlers' design to further anthropomorphize the animals and engage the target-language audience in a different experience (e.g. more empathic to the dog), as opposed to replicating the experience of the source-language audience (McClarty 2014). As such, creative subtitling does not seem to care much about "representational accuracy, fidelity, and authenticity"; instead, "affectivity, subjectivity and social engagement are the main drivers" (Guillot 2019: 36; see also Pérez-González 2014).

The previous analyses echo what Díaz Cintas and Remael (2021) remind us: "Quantity and quality are hardly the same" (147). Thus, for the systematic analysis of graphic elements in creative subtitling, it is important to (a) go beyond the face-value comparison of modal resources between the source and target subtitles and (b) map the modal repertoires onto the (in)variations of metafunctional meanings.

5.2 Opportunities and challenges of a corpus-based approach

Different from previous case-based studies on creative subtitling, the current study takes into account a wider range of graphic elements and the semiotic meanings they encode (see Figure 2). Although this study has limited itself to one YouTube channel and a relatively small corpus, as Soffritti (2019) rightly emphasizes, small corpora can also make valuable contributions "as an empirical basis" and "a starting point" for future research (345). As

demonstrated in this study, a corpus-based approach to multimodality and creative subtitling affords promising opportunities in at least two aspects. First, the approach enables researchers to better discern and disentangle multimodal complexity that undergirds creative subtitling. As illustrated in the previous section, the approach calls for associating the modal shifts with the (in)variations of metafunctional meanings. This epistemological lens reveals the multimodal strategies and the underlying agendas. As discussed previously, the patterned uses of emojis suggested that the subtitlers might have prioritized affectivity over accuracy and audience engagement over textual fidelity. A fruitful avenue will be to expand the current scope of *parallel* multimodal corpora to include *comparable* multimodal corpora for triangulation (Bernardini and Ferraresi 2022; Malamatidou 2018; Zanettin 2012). For instance, it will be interesting to include the English multimodal subtitles and Chinese multimodal subtitles that are untranslated and intended for native-language speakers on streaming platforms. Pitting these comparable corpora against the parallel multimodal corpora can further reveal how the semiotic landscapes intersect, as orchestrated by the subtitlers and the content creators.

A second opportunity afforded by the corpus-based approach is to draw on naturally occurring data to inform the theorization of co-contextualization of multimodal resources in creative subtitling. For two decades, scholars have called for due attention to the modal interplay in subtitling, particularly the visual-verbal interplay (Gambier 2003, 2013; Guillot 2019). Previous studies have explored how subtitles are shaped by the visual information in the preceding, accompanying, and following scenes (Chen 2019). In this way, subtitles and images have been treated as two discrete modal entities. While this treatment seems acceptable for conventional subtitling, it becomes problematic for creative

subtitling that draws on more than the language mode. As illustrated in this study, the subtitlers were liberal with the use of graphic/visual elements in both source and target subtitles. Conventional subtitling tends to take visual elements as a mere multimodal background (Tuominen et al. 2018). This text-context treatment arguably centers the language mode and relegates the visual mode to a secondary or even peripheral position. However, creative subtitling transcends the modal boundaries between the verbal and the visual and challenges the dichotomy between the text and the context. With the integration of graphic elements and subtitle texts, all modes maintain co-contextualizing relations with one another in creative subtitling. As in this study, all modal resources make contribution to the semiotic meaning-making (Bateman 2008), such as emojis to mark affectivity, soundtracks and brackets to identify speakers, images and subtitle texts to represent the narrative. As such, a corpus-based approach can systematically attend to and account for the varieties, combinations, and interactions of modal resources related to creative subtitling. In addition, when we trace how the co-contextualizing relations shift from the source to the target subtitles, the comparisons can reveal what and how multimodal resources are foregrounded and backgrounded in response to the rhetorical exigence prompted by translation. As Matthiessen (2021: 524) aptly points out, translation choices "are *probabilistic* in nature…[and] need to be investigated in terms of corpora" (original emphasis). Thus, the patterned shifts of co-contextualization relations observed from the corpus data can shed light on the frequency, range, and aptness of multimodal choices affecting and affected by creative subtitling.

Despite the opportunities outlined previously, the corpus-based approach to the graphic elements in creative subtitling is not without its challenges. The first challenge is

to deal with the fluidity of semiotic meanings that potentially vary from one community to another. While it is tempting to assume that the use of emojis in the current study minimizes misunderstanding, Díaz Cintas and Remael (2021: 65) caution that "images do not necessarily carry universal meaning since they are often culturally determined." Similarly, the social semiotic approach to multimodality emphasizes that "meanings are socially made, socially agreed and consequently socially and culturally specific" (Kress 2010: 88). Previous research has indeed shown that people with different cultural backgrounds may interpret emojis differently (Sun, Lasser and Lee 2022). In this study, for the purpose of corpus annotation, Emojipedia was consulted to decide whether an emoji was used in a contextually congruent manner. When none of the typical emotions listed for an emoji matched its context of use, it was coded as divergent-ideational. In fact, 17 instances out of the 48 divergent-ideational meanings (or 35.4%) were enacted due to the "Face with Steam from Noise" emoji (see Example 2 in Figure 3). This over-representation might suggest that the subtitlers intended the emoji to mean differently from the possible emotional states normally recognized by the English-speaking viewers. Thus, based on the over- or under-represented frequency counts, the corpus-based approach can point to mismatched semiotic meanings between the source and the target communities. However, it is difficult to go beyond this descriptivism and ascertain whether the viewers and subtitlers experience the fluid meanings as such, which provides an appropriate segue for the discussion of the next challenge.

The second potential limitation of the corpus-based approach is the difficulty to extrapolate corpus analysts' interpretation to real-life audience reception. This is perhaps an inherent issue for AVT studies drawing on corpora as data sources (Wu and Chen 2022).

The previous paragraph highlights the variability of semiotic meanings; and it should also be stressed that audiences are by no means static either. In the words of Guillot (2019: 35), the "viewing and processing habits [are] in a state of flux as audiences get more adept at processing text on screen in semiotically complex contexts in the age of online multimedia." As audience reception is an essential part of AVT studies, corpus analysts' interpretation requires investigator and methodological triangulation (Malamatidou 2018) to generate comprehensive and robust insights into audience perceptions and preferences. This study adopted investigator triangulation by having the coupled pairs double-coded to ensure the inter-subjectivity of the coding. For future studies, methodological triangulation will be a promising avenue. For instance, an experiment can be conducted on two groups of viewers, exposed to either conventional subtitling or creative subtitling. Eye-tracking data, perception questionnaires, and emotional ratings will shed light on how they differ in cognition, perception, and emotion. Another experiment can be conducted on viewers with different cultural backgrounds but exposed to the same set of creative subtitles to ascertain whether they perceive the metafunctional meanings to be similar or otherwise. The findings will substantiate or refine the claims of metafunctional meanings enacted by the modal shifts as observed in the current corpus-based study. Finally, it will be interesting to interview the subtitlers and the viewers to see how the meaning-making and sense-making converge or diverge in the production and reception of creative subtitling. When triangulated with other research methods, the corpus-based approach will not be limited to the *etic* examination of co-contextualization of semiotic resources bundled in subtitle texts, but will be complemented with *emic* insights from agents in socio-semiotic milieus.

6. Conclusion

This study adopted a corpus-based approach to the graphic elements in creative subtitling on a YouTube channel. It was found that the source and target subtitles drew on different graphic elements as multimodal resources. Despite the general trend of a reduced number of modes in the target subtitles, the metafunctional meanings tended to be complementary or similar to the source subtitles. These corpus findings shed light on the subtitlers' semiotic considerations and their prioritized agendas that motivated the shifts of graphic elements in the target subtitles. Overall, this study points to the opportunities afforded by the corpus-based approach to unraveling the multimodal complexity and theorizing the co-contextualization relations of multimodal resources in creative subtitling. These come with the caveat that future corpus-based studies should be mindful of the fluidity of semiotic meanings and the necessity of investigator and methodological triangulation. When corpus findings are enriched by inter-subjective and etic-emic insights, they will afford a fuller, more sophisticated account of creative subtitling.

References

Primary sources

Video links for the examples in Figure 3.

Example 1: https://www.youtube.com/watch?v=woiY8H0wLJ4 (2:44)

Example 2: https://www.youtube.com/watch?v=Y2rChaREC5g (1:20)

Example 3: https://www.youtube.com/watch?v=G_rOvmYpwZ8 (0:12)

Example 4: https://www.youtube.com/watch?v=mu6pmo4e3ek (0:05)

Secondary sources

Agresti, Alan. 2013. *Categorical Data Analysis*. Hoboken: John Wiley & Sons.

Baños, Rocío, Silvia Bruti, and Serenella Zanotti. 2013. "Corpus linguistics and Audiovisual Translation: in search of an integrated approach." *Perspectives* 21 (4): 483–490.

Bassnett, Susan, Lawrence Venuti, Jan Pedersen, and Ivana Hostová. 2022. "Translation and Creativity in the 21st Century." *World Literature Studies* 14 (1): 3–17.

Bateman, John. 2008. *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. London: Palgrave Macmillan.

Bateman, John A. 2014. "Using multimodal corpora for empirical research." In *The Routledge Handbook of Multimodal Analysis*, edited by Carey Jewitt, 238–252. London: Routledge.

Beckman, Rachel. 2008. "An Out-of-character Role for Subtitles." Washington Post. https://www.washingtonpost.com/wpdyn/content/article/2008/11/14/AR2008111400 700.html. Assessed on 20 August, 2023.

Bernardini, Silvia, and Adriano Ferraresi. 2022. "Corpus Linguistics." In *The Routledge Handbook of Translation and Methodology*, edited by Federico Zanettin and Christopher Rundle, 207–222. London: Routledge.

Brookes, Gavin, and Tony McEnery. 2020. "Corpus Linguistics." In *The Routledge Handbook of English Language and Digital Humanities*, edited by Svenja Adolphs and Dawn Knight, 378–404. London: Routledge.

Bruti, Silvia. 2020. "Corpus Approaches and Audiovisual Translation." In *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*, edited by Łukasz Bogucki and Mikołaj Deckert, 381–395. Cham: Palgrave Macmillan.

Bywood, Lindsay. 2019. "Testing The Retranslation Hypothesis for Audiovisual Translation: The Films of Volker Schlöndorff Subtitled into English." *Perspectives* 27 (6): 815–832.

Chen, Yuping. 2019. *Translating Film Subtitles into Chinese: A Multimodal Study*. Singapore: Springer.

Cohn, Neil. 2013. "Beyond speech balloons and thought bubbles: The integration of text and image." *Semiotica* 197: 35–63.

Dahne, Annie, and Roberta Piazza. 2020. "'Subtitles Have to Become My Ears Not My Eyes' Pragmatic-stylistic Choices Behind Closed Captions for The Deaf and Hard of Hearing: The Example of *Breaking Bad*." In *Telecinematic Stylistics*, edited by Christian Hoffmann and Monika Kirner-Ludwig, 285–306. London: Bloomsbury.

Danesi, Marcel. 2017. *The Semiotics of Emoji: The Rise of Visual Language in the Age of the Internet*. London: Bloomsbury.

De Linde, Zoé, and Neil Kay. 1999. "Processing Subtitles and Film Images: Hearing vs Deaf Viewers." *The Translator* 5(1): 45–60.

Díaz Cintas, Jorge. 2018. "'Subtitling's a Carnival': New Practices in Cyberspace." *The Journal of Specialised Translation* 30: 127–149.

Díaz Cintas, Jorge, and Aline Remael. 2021. *Subtitling: Concepts and Practices*. London: Routledge.

Dwyer, Tessa. 2017. *Speaking in Subtitles: Revaluing Screen Translation*. Edinburgh: Edinburgh University Press.

Field, Andy. 2018. *Discovering Statistics Using IBM SPSS Statistics (5th edition)*. Los Angeles: Sage.

Foerster, Anna. 2010. "Towards a Creative Approach in Subtitling: A Case Study." In *New Insights into Audiovisual Translation and Media Accessibility*, edited by Jorge Díaz Cintas, Anna Matamala, and Josélia Neves, 81–98. Amsterdam: Rodopi.

Fox, Wendy. 2016. "Integrated Titles: An Improved Viewing Experience." In *Eyetracking and Applied Linguistics*, edited by Silvia Hansen-Schirra and Sambor Grucza, 5–29. Berlin: Language Science Press.

Fox, Wendy. 2018. *Can integrated titles improve the viewing experience?: Investigating the impact of subtitling on the reception and enjoyment of film using eye tracking and questionnaire data*. Berlin: Language Science Press.

Gambier, Yves. 2003. "Introduction: Screen transadaptation: Perception and reception. Special issue on screen translation." *The Translator* 9 (2): 171–190.

Gambier, Yves. 2013. "The Position of Audiovisual Translation Studies." In *The Routledge Handbook of Translation Studies*, edited by Carmen Millán and Francesca Bartrina, 45–59. London: Routledge.

Guillot, Marie-Noëlle. 2019. "Subtitling on the Cusp of its Futures." In *The Routledge Handbook of Audiovisual Translation*, edited by Luis Pérez-González, 31–47. London: Routledge.

Halliday, Michael Alexander Kirkwood. 1978. *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Halliday, Michael Alexander Kirkwood. 2001. "Towards a Theory of Good Translation."
In *Exploring Translation and Multilingual Text Production: Beyond Content*, edited
by Erich Steiner and Colin Yallop, 13–18. Berlin: Mouton de Gruyter.

Halliday, Michael Alexander Kirkwood, and Christian M.I.M Matthiessen. 2014.
Halliday's Introduction to Functional Grammar. London: Routledge.

Hart, Christopher. 2020. "Multimodal Discourse Analysis." In *Researching Discourse: A
Student Guide*, edited by Christopher Hart, 143–179. London: Routledge.

Jewitt, Carey, Jeff Bezemer, and Kay O'Halloran. 2016. *Introducing Multimodality*.
London: Routledge.

Kress, Gunther. 2010. *Multimodality: A Social Semiotic Approach to Contemporary
Communication*. London: Routledge.

Kress, Gunther. 2014. "What is a mode?" In *The Routledge Handbook of Multimodal
Analysis*, 2nd edn, ed.by Carey Jewitt, 60–75. London: Routledge.

Kress, Gunther. 2020. "Transposing Meaning: Translation in a Multimodal Semiotic
Landscape." In *Translation and Multimodality: Beyond Words*, edited by Monica
Boria, Ángeles Carreres, María Noriega- Sánchez and Marcus Tomalin, 24–48.
London: Routledge.

Kress, Gunther, and Theo Van Leeuwen. 2006. *Reading Images: The Grammar of Visual
Design*. 2nd edn. London: Routledge.

Kruger, Jan-Louis. 2021. "Translation, Multimodality and Cognition." In *The Routledge
Handbook of Translation and Cognition*, edited by Fabio Alves and Arnt Lykke
Jakobsen, 433–444. London: Routledge.

Levshina, Natalia. 2017. "Online Film Subtitles as A Corpus: An N-Gram Approach." *Corpora* 12 (3): 311–338.

Logi, Lorenzo, and Michele Zappavigna. 2021. "A Social Semiotic Perspective on Emoji: How Emoji and Language Interact to Make Meaning in Digital Messages." *New Media & Society*. DOI: 14614448211032965.

Malamatidou, Sofia. 2018. *Corpus Triangulation: Combining Data and Methods in Corpus-Based Translation Studies*. London: Routledge.

Matthiessen, Christian M. I. M. 2021. Translation, multilingual text production and cognition viewed in terms of systemic functional linguistics. In *The Routledge Handbook of Translation and Cognition*, edited by Fabio Alves and Arnt Lykke Jakobsen, 517–544. London: Routledge.

McClarty, Rebecca. 2012. "Towards A Multidisciplinary Approach in Creative Subtitling." *MonTI: Monografías de traducción e interpretación* 4: 133–153.

McClarty, Rebecca. 2014. "In Support of Creative Subtitling: Contemporary Context and Theoretical Framework." *Perspectives* 22 (4): 592–606.

McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

Miles, Matthew B., and A. Michael Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. 2 edn. Thousand Oaks, CA: Sage.

Mubenga, Kajingulu. 2009. "Towards a Multimodal Pragmatic Analysis of Film Discourse in Audiovisual Translation." *Meta* 54 (3): 466–484.

Nornes, Abé Mark. 1999. "For an Abusive Subtitling." *Film Quarterly* 52 (3): 17–34.

Pavesi, Maria. 2019. "Corpus-based Audiovisual Translation Studies: Ample Room for Development." In *The Routledge Handbook of Audiovisual Translation*, edited by Luis Pérez-González, 315–333. Abingdon: Routledge.

Pavesi, Maria. 2022. "Corpora and The Language of Films: Exploring Dialogue in English and Italian." In *The Routledge Handbook of Corpus Linguistics* (second edition), edited by Anne O'Keeffe and Michael J. McCarthy, 547–561. Abingdon: Routledge.

Pedersen, Jan. 2011. *Subtitling Norms for Television: An Exploration Focussing on Extralinguistic Cultural References*. Amsterdam: John Benjamins.

Perego, Elisa. 2019. "Into the language of museum audio descriptions: a corpus-based study." *Perspectives* 27 (3): 333–349.

Pérez-González, Luis. 2007. "Appraising Dubbed Conversation: Systemic Functional Insights into the Construal of Naturalness in Translated Film Dialogue." *The Translator* 13 (1): 1–38.

Pérez-González, Luis. 2014. *Audiovisual Translation: Theories, Methods and Issues*. London: Routledge.

Pinnow, Rachel J. 2015. "A Patriot is Respectful: (Re-)Examining the Architecture of Ideology in Educational Contexts." In *International Handbook of Semiotics*, edited by Peter Pericles Trifonas, 1249–1269. New York: Springer.

Pinto, Sara Ramos, and Aishah Mubaraki. 2020. "Multimodal Corpus Analysis of Subtitling: The Case of Non-Standard Varieties." *Target* 32 (3): 389–419.

Remael, Aline, and Nina Reviers. 2020. "Media Accessibility and Accessible Design." In *The Routledge Handbook of Translation and Technology*, edited by Minako O'Hagan, 482–497. London: Routledge.

Reviers, Nina. 2018. "Studying the Language of Dutch Audio Description: An Example of a Corpus-Based Analysis." *Translation and Translanguaging in Multilingual Contexts* 4: 178–202.

Romero-Fresco, Pablo. 2009. "Naturalness in the Spanish Dubbing Language: A Case of Not-so-close Friends", *Meta* 54 (1): 49–72.

Soffritti, Marcello. 2019. "Multimodal Corpora in Audiovisual Translation Studies." In *The Routledge Handbook of Audiovisual Translation*, edited by Luis Pérez-González, 334–349. London: Routledge.

Sun, Juhyung, Sarah Lasser, and Sun Kyong Lee. 2022. "Understanding Emojis: Cultural Influences in Interpretation and Choice of Emojis." *Journal of International and Intercultural Communication*. 1–20.

Taylor, Christopher. 2003. "Multimodal Transcription in the Analysis, Translation and Subtitling of Italian Films." *The Translator* 9 (2): 191–205.

Taylor, Christopher. 2013. "Multimodality and Audiovisual Translation." In *Handbook Of Translation Studies* (Volume 4), edited by Yves Gambier and Doorslaer, Luc van, 98–104. Amsterdam: John Benjamins.

Taylor, Christopher. 2016. "The Multimodal Approach in Audiovisual Translation." *Target* 28 (2): 222–236.

Taylor, Christopher. 2020. "Multimodality and Intersemiotic Translation." In *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*, edited by Łukasz Bogucki and Mikołaj Deckert, 83–99. Cham: Palgrave Macmillan.

Thompson, Paul. 2022. "Building A Specialised Audiovisual Corpus." In *The Routledge Handbook of Corpus Linguistics (2nd edition)*, edited by Anne O'Keeffe and Michael J. McCarthy, 75–88. London: Routledge.

Tirkkonen-Condit, Sonja, and Jukka Mäkisalo. 2007. "Cohesion in Subtitles: A Corpus-based Study." *Across Languages and Cultures* 8 (2): 221–230.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Toury, Gideon. 2012. *Descriptive Translation Studies and Beyond* (Revised edition). Amsterdam: John Benjamins.

Tuominen, Tiina, Catalina Jiménez Hurtado, and Anne Ketola. 2018. "Why Methods Matter: Approaching Multimodality in Translation Research." *Linguistica Antverpiensia: New Series–Themes in Translation Studies* 17: 1–21.

Valentini, Cristina. 2007. "Developing AVT Corpora for a Quantitative Approach to Language Transfer in Cinematic Products." In *Selected Papers of the CETRA Research Seminar in Translation Studies 2006*, edited by Francis Mus, 1–18. Leuven: CETRA.

Wu, Zhiwei, and Zhuojia Chen. 2022. "Towards a corpus-driven approach to audiovisual translation (AVT) reception: A case study of YouTube viewer comments." The Journal of Specialised Translation 38: 128–154.

Xavier, Catarina. 2022. "A Three-Layered Typology for the Subtitling of Taboo: A Corpus-Based Proposal of Methods, Strategies, and Techniques." *Babel* 68 (4): 586–609.

Zanettin, Federico. 2012. *Translation-Driven Corpora: Corpus Resources for Descriptive and Applied Translation Studies*. Manchester: St. Jerome.