

REPETITION AND ADDITION AS COHERENT DEVICES IN AUDIO DESCRIPTION OF COMICS: A CASE STUDY OF *X: BIG BAD*

Chunwei Liu*
University of Glasgow

ABSTRACT: Audio Description (AD) translation is an important translational practice that promotes the accessibility of visual texts for the visually impaired audience. Although AD translation has been developed to be a regular practice in performing arts and entertainment, its potential has not been fully explored for visual-verbal sequential arts such as comics to meet the demand of the audience. To fill in such a gap, this article performs a case study on a commercial AD translation of an action comic series *X: Big Bad* to explore translation strategies as norms that specifically facilitate achieving coherence and are unique to the cross-modal translation of comics, aiming at providing theorised guidelines for future practices. Based on visual language and narratology analysis, three types of visible interventions in AD comics translation are identified, challenging the notion of neutrality in AD and arguing that reasonable additions and adaptations can facilitate the target text's accessibility.

KEYWORDS: Multimodal translation; Audio description for comics; Accessible aids for visually impaired readers; Coherence

1. Introduction

Audio description (AD) refers to audio tracks that use verbal descriptions and other types of audible techniques to make visually based materials accessible to visually impaired readers and audiences (Kleege, 2016, p. 89). It is one of the key methods of promoting the accessibility of visual or verbal-visual materials for blind or visually impaired readers. Its popularity and necessity have been widely accepted in audible-visual practises, and relatively matured methodologies, regulations and production guidance in relevant fields have been developed to facilitate practitioners and researchers (Packer et al., 2015). However, compared to the thriving development of AD translation for movies, TV shows and theatres (Kruger and Orero, 2010), AD translation for comics remains under-explored, both practically and academically, despite a vast potential readership.

Guy Hasson, an advocate for accessible comics and the founder of an AD comic production company called Comics Empower, expresses concern about the lack of resources for visually impaired comic readers:

“There are literally tens of thousands of kids, teens, and adults out there who would love to read the same comic books their sighted friends are reading and to talk to them about it” (Hasson cited in Kingett, 2016).

In response to this demand, several important AD translation projects have been carried out using different approaches. Hasson focuses on developing AD scripts unique to comics

*2431662C@student.gla.ac.uk

as a visual genre, and his translational approach recreates a visual-like reading experience for the target audience (ibid.). Although his business has now unfortunately ceased to operate, the comic-based approach has been taken up by Rachel Osolen and Leah Brochu in their project of audio-describing *The Walking Dead* (2018), accompanied by a comprehensive translation report (2020). As well as making an effort to rationalise and standardise scriptwriting for comics AD (Osolen and Brochu, 2020, p. 110), the production team also argues for the importance of objectivity. As a part of their practical research project, Osolen and Brochu reviewed commonly accepted AD guides and found a shared emphasis on objectivity, ‘as it is not the job of the describer to give their interpretation of the situation’ (ibid.).

Brandon Christopher, on the other hand, argues that visual-to-non-visual translation is ‘bound to result in a certain amount of signal loss’ (2018), which means that adaptation is almost inevitable in audio comic translation. Similar observations about adaptation can be found in practitioners’ reports, which note features such as the tolerance of repetitions (McGee-Tubb, 2017), consideration of colouring and background (ibid.), and an emphasis on key details and time/spatial references (Hasson cited in Osolen and Brochu, 2020, p. 111).¹ Thus, the multimodal nature of both STs and TTs determines the need for interpretation of the TT. As Christopher Taylor observes, with regard to the AD translation of narrations: ‘[a]ll translations involve some level of reformulation from a source text’ (2016, p. 224). Considering the prevalence of the notion that the AD translator should remain invisible, the need for adaptation raises further questions: if a key detail or reference is not directly depicted in one panel but either implied in relevant visual elements or provided in contextual panels, to what extent are interpretations and adaptations acceptable for ensuring coherence in the TT? What translation strategies and aural devices can be applied to build cohesion and consistency?

As the level of interpretation is far less theorized in AD comics research—with relevant practical experience only vaguely referred to as ‘striking a balance’ (McGee-Tubb, 2017)—it is worth mentioning another approach, which aims to create a ‘movie experience’. This, according to Hasson (cited in Kingett, 2016), is a more common target for larger commercial productions. For example, an audio version of the first issue of *DareDevil* released by Marvel Comics in 2011 (unfortunately no longer accessible on the official website) derives directly from the original script of the source text (ST) rather than the visual text itself to create seamless narrative that is almost similar to an audio drama, while ignoring the panel divisions and gutters of the typical comic. Given the significant discrepancies that this approach produces in relation to the visual source text, it has led to frequent criticism.

¹ Unfortunately, most of the first-hand materials related to Hasson’s work, such as his company, translations and podcasts gradually become unavailable after his business stopped operating in 2017 (<https://twitter.com/ComicsEmpower/status/972369143516327936>). Consequently, the quotations about his work in this article have been extracted from secondary sources.

On the other hand, GraphicAudio (an AD comic and e-book company that works more closely with verbal-visual STs and produces descriptions by translating and adapting popular comics comparatively, in large quantities and with stable quality control) applies cohesive devices and aural mechanisms such as background music and sound effects to create coherently organised and commercially matured TTs (their slogan is ‘a movie in your mind’), even if these may appear inaccurate at times. In general, although the company has never systematically explained their working procedures, the presence of intermodal translation is more noticeable in GraphicAudio’s TTs, while their productions also exhibit a relatively high level of coherence.

This article focuses on coherence and cohesion as a translation goal of AD comics and attempts to generalise from one successful production in order to identify potential norms that could be used by the genre in future. It compares the information in the ST and TT in order to identify the strategies used in the audiodescription and analyse how they contribute to achieving coherence across the target text as a whole. It concludes that the interpretation of visual elements and their interaction with verbal texts is unavoidable for AD comic translation, and to achieve a coherent target text, it is reasonable to apply relatively visible strategies such as highlighting specific objects or adding additional narrative perspectives and transitioning movements.

2. The Case Study

X: Big Bad, initially released by Dark Horse (Swierczynski et al., 2014), is an action comic featuring a vigilante called X and a freelance reporter Leigh, depicting their quest for revenge against Berkshire, a local oligarch. It is a typical superhero comic with an intensive narrative arc and characters with distinguishable visual designs, which employs complex techniques to express motion and time duration through static means. However, its visuals occasionally lack some of the features that we have come to expect from an action comic, such as consistent recurring details or strong visual connections between panels – the result of conscious artistic choices, one supposes. These raise some challenges for audiodescription, as we shall see.

As for the TT, this was produced by GraphicAudio (Messner, 2021) and officially published in collaboration with the ST publisher Dark Horse, which implies a certain level of quality control, authority, and acceptance. In what follows, the 134-page (covers and section covers included) comic and 1:56:06 audio track are manually compared in order to identify noticeable discrepancies in content, specifically focusing on potential additions and adaptations.

As for the analytical framework used, this relies extensively on visual language and narrative theories, including McCloud’s *Understanding Comics* (1994) and *Making Comics* (2006) for visual and formatting norms in comics, Groensteen’s theory of narratology in comics (2013), Gavalier’s categorisation of non-narrative panel transitions (2017), and Cohn’s visual language grammar for narrative cohesion (2013).

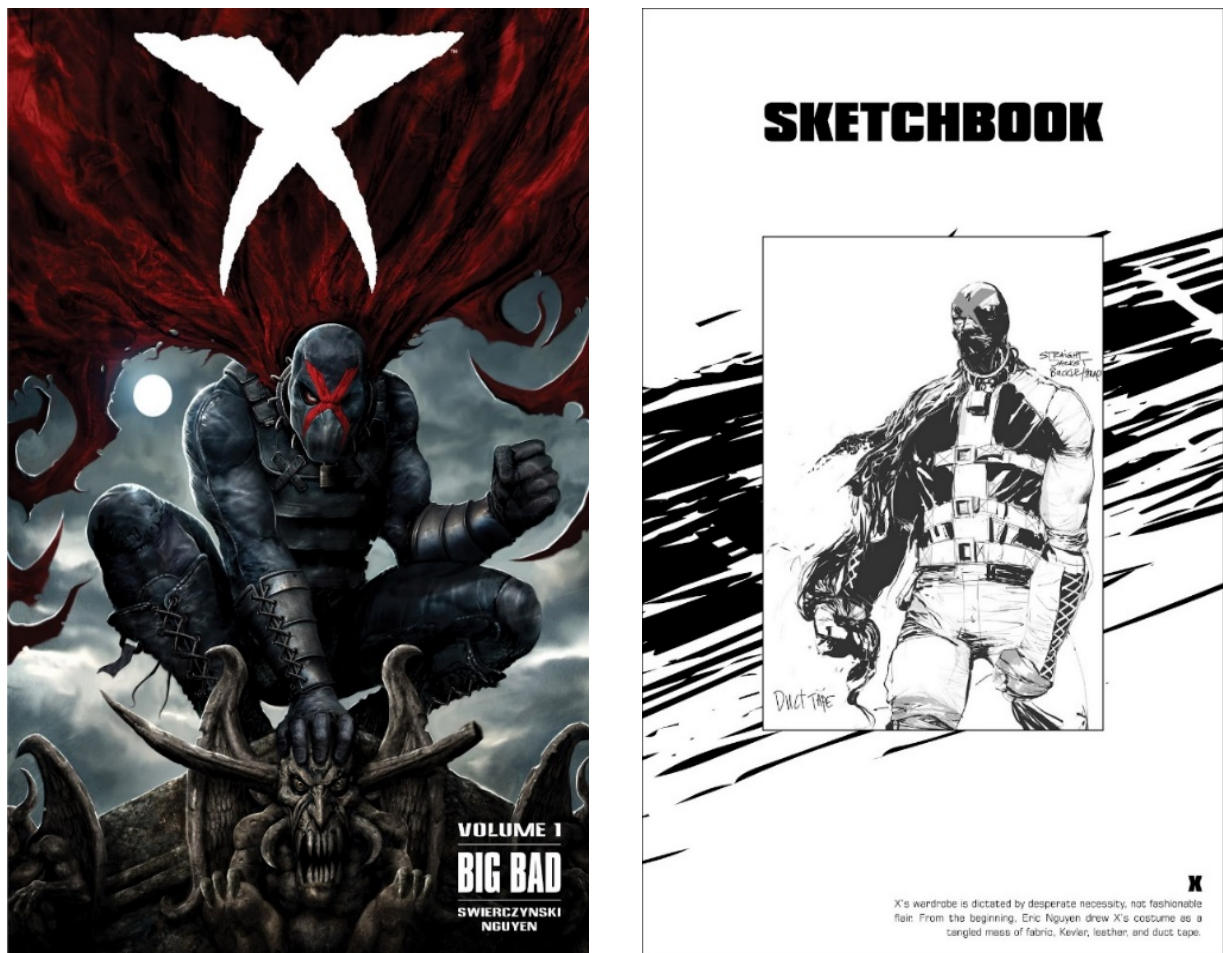


Figure 1. X's costume design on the issues' cover and in the artist's notes (Swierczynski et al., 2014, cover page + p. 126)

3. Findings and Evaluation

3.1 In-panel challenges: background information and recurring items

A common practice for developing narrative threads in comics is the use of recurring items (Groensteen and Miller, 2013, p. 12). In the source text, the repetition of visual representations is frequently used for purposes such as character-building or narrative development. For instance, specific visual elements, such as the lock or cape on X's superhero outfit are consistently and meticulously portrayed. These recurring visuals serve multiple purposes within the narrative, such as advancing the plot or building suspense around X's identity (see Figure 1). On the front page, X's cape takes up around 1/3 of the composition in a dramatic and dynamic manner, promoting the protagonist's signature attire as identifiable, visually attractive gear with his name branded at the centre. Similarly, the artist's note in the second image also confirms the importance of this 'tangled mass of fabric' (Swierczynski et al., 2014, p. 126), as it implies the protagonist's personality traits of being practical and realistic. Failure to acknowledge the significance of repetition during translation may result in a less coherent and consequently less accessible description.

To address this issue, the audiodescription of *X: Big Bad* employs two main strategies: intentionally increasing the frequency of recurring visuals, and incorporating contextual background information to provide additional context.

As mentioned earlier, X's vigilante costume is a key symbol of his strength and anonymity, as, in superhero comics, 'the superhero costume is a kind of mask, allowing him to masquerade within a constructed identity' (Brownie and Graydon, 2016, p. 27). The cape in particular contributes to the plot development in X's fight with Berkshire and the police on pages 47 to 55. As the first scene of this narrative arc, the first panel on Page 47 shows X standing on the rooftop of a skyscraper (Figure 2: Swierczynski et al., 2014). By comparing this panel with the cover art (Swierczynski et al., 2014, cover) and the character design sketch (ibid., p. 126) in Figure 1, the cape as part of his costume stays consistent from the initial draft to the final visualised production.

As the artist's note suggests, 'X's wardrobe is dictated by desperate necessity' (ibid.), as exemplified by the scene in Panel 2, page 54, where X utilizes his cape to shield himself



Figure 2. Page 47 as an example for consistency in describing key items (Swierczynski et al., 2014)

from an explosion (Swierczynski et al., 2014). In this panel (reproduced in Figure 3), X covers his head and body with the cape before he ignites a fuse. Because the subsequent panel

depicts him being propelled by an explosion, yet overall unharmed, a subtle cause and effect relationship is created between the protection offered by the cape and X's survival. In this regard, X's cape as a recurring item is both an inseparable part of the character's image and a necessary plot device in the narrative arc.

To transmit the importance of this gear in audiodescription, the TT repeats the term 'cape' almost to the point of appearing repetitive and cumbersome. For Panel 1 (page 47), the audiodescription reports 'the wind fluttered X's tattered cape' (Messner, 2021, 39:46-39:49), and only around ten seconds later, for Panel 2, his costume is again highlighted with 'the torn cape billowed out behind him (Messner, 2021, 40:00-40:01). Both 'tattered'² and 'torn'³, two synonyms for 'ragged', helps the TT's audience to visualise the condition of X's cape and develop awareness of his character image, which prepares for the cape's utility on page 54. Another effort the TT makes is to introduce additional information about the cape's function, which directs the audience's attention beyond its symbolic significance. Referring to the costume again as 'tattered and torn', the TT describes Panel 2 (page 54) as follows:

In the fire tower, X crouched over the wire he's pulled out. He covered himself with his cape. Though tattered and torn, the cape was reinforced with Kavalier and would offer him a little protection. He struck a match *male voice groaning*, lit the end of the wire and ran (Messner, 2021, 44:09-44:24).

In this extract, not only is the cape and its condition reiterated using the exact same wording as employed approximately four minutes prior, but a detailed explanation is also given as to why this cape would help X survive the explosion. Since the material of the cape cannot be read either from this panel, from previous or following story arcs or from the character designs in the appendix, this background information about the cape's details is clearly a TT addition. It draws the audience's attention to X's protective equipment and prepares them for his safe escape, successfully achieved in the description of Panel 3: 'a caped figure plunged through the air' (Messner, 2021, 44: 40-44:43). Again, coherence is developed by mentioning the cape to make sure this scene does not sound abrupt and unexpected to the audience.

² <https://www.merriam-webster.com/dictionary/tattered> (Accessed: 1 April 2023).

³ <https://www.merriam-webster.com/dictionary/torn> (Accessed: 1 April 2023).



Figure 3. Page 54 as an example for consistency in describing key items (Swierczynski et al., 2014)

Other cases of recurrence can be found throughout this excerpt, even in less dramatic scenes. For example, ‘dumpsters’ is repeated three times within 20 seconds (Messner, 2021, 43:52, 43:57, 44:06) to help the audience imagine the relation between Leigh and the back alley where she accidentally rescues X. The ‘heavily armed’ police officer’s ‘walkie-talkies’ and ‘machine guns’ also appear multiple times (Messner, 2021, 41:12-43: 12) to constantly remind the audience that the antagonist threatens X with force. X’s tactic belt is also described on two occasions, the first time for page 47 Panel 2 (Swierczynski et al., 2014) and the second for page 53 Panel 1 (ibid.), with both descriptions explaining where X’s weapons come from (Messner, 2021, 39:55, 43:18). For the later repetition, the panel it describes is a close-up shot of X’s hand so the belt is clearly not a part of the ST visual; however, in the next panel, X’s crouched position means that his belt cannot be seen by the ST’s readers. The TT addition, while not strictly corresponding to the visuals, maintains the coherence established by the previous description of X’s attire and reinforces the image of a heavily armed superhero.

Overall, it is evident that the TT displays a notable tendency to repeat items by consistently employing the same wording, particularly when the item's significance extends to subsequent narrative developments. The TT also tends to add information about the following panels to develop a semantic correlation between the descriptions of one panel to another and thus build a strong context. Under such circumstances, a reasonable level of interpretation and addition is clearly deemed acceptable. As stated by Georgina Kleege (2016, p. 96), without interpretation of this kind, 'audio description often withholds information in a way that can draw undue attention to the absent information'.

3.2 Multiple non-narrative panels: Constructing the narrative perspective

The narrative perspective is the angle from which a storyline is presented, a feature shared by both comics and ADs. Perspectives can be primarily categorized into first-, second- and third-person; however, comics are unique in that the perspectives they offer can be further categorised into those of the *narrator* who controls the storytelling (who may be invisible), the *reciter* who offers the visual perspective, and the *monstrator* who verbally narrates the story (Groensteen, 2013, pp. 79-86). The latter category can complement the basic categorisation for a more precise description of the perspective (for example, the first-person narrator or the third-person narrator).

In the visuals of this ST, there is a tendency to alternate between a third-person omniscient narrator and the third-person reciters, which include characters such as Leigh, X, and the villains, along with occasional first-person reciters in fighting scenes. The verbal textboxes, on the other hand, are presented by both the first-person monstrator and the third-person narrator, which complicates the overall perspective and poses challenges for audiodescription. For instance, a visual drawn from the third-person narrator's perspective may be accompanied by a first-person narration or train of thought, presented in square boxes. The multimodality of these perspectives contributes to the cohesion of the whole, and when such verbal hints are subtle or even missing, the textual narrations intervene and build the connection between one panel and another. In the TT, however, the lack of visuals makes it impossible to provide scene changes and verbal ties at the same time. Transitioning between independent scenes or from current events to flashbacks needs to be supported with additional contextual information, in many cases, reference to perspectives.

Even in narrative comics, a small number of panels may be connected through non-narrative relations or arranged in an unordered manner (Gavaler, 2017, pp. 19-23), raising challenges for clarity and cohesion in audiodescription. Figure 4 depicts how Leigh collects information from newspapers, providing an example of weakly-linked visuals which have to be represented using sign-postings (especially movements) in AD to reinforce a coherent perspective of a specific character. The initial four panels adopt a third-person perspective to show Leigh's discoveries, but, without the verbal texts building cohesion, they cannot form a coherent narrative arc. Depicting respectively a party scene, an election campaign, a construction site and a headshot of a man drinking wine (Swierczynski et al., 2014, p. 64), Panels 1-4 exemplify what Gavaler classifies as a representational arrangement, meaning



Figure 4. Page 64 as an example for adding narrative perspectives (Swierczynski et al., 2014)

that their connections are achieved by their logical relationships instead of through sequential or narrative coherence (ibid.). Moreover, the four panels are not consistent as regards the presence of recurring images: although Panels 1, 2 and 4 all feature Berkshire (referred to as 'Mr. Juicy' in the texts), Panel 3 does not, which breaks the coherence.⁴ And even though these panels all have a different colour palette to the main narration arc (mainly greens, blues, purples, and greys versus a variety of reds and yellows, which gives them a colder tone), implying that they represent an interpolation that is independent of the major storyline,⁵ it does not give further translatable information to help with the description. Thus, if the panels were to be described as they are without intervention, the logical progression in the TT would appear arbitrary and be difficult to understand.

⁴ See Steiner (2004) and Eisner (2008) on the significance of recurring images in maintaining coherence in comics.

⁵ See Groensteen (2013, p. 153) on the use of colour variations to mark different sections and paces.

To address this lack of connection, the verbal texts (enclosed in square boxes) in Panels 1-4 do suggest a logical context for the arrangement of the visuals, providing background information about a local political and economic oligarch. Here, the use of square boxes⁶ means that the texts are not actual dialogues happening between two characters but in fact monologues, more specifically a blog article written by Leigh (the female figure in Panel 5) as the first-person monstrator.


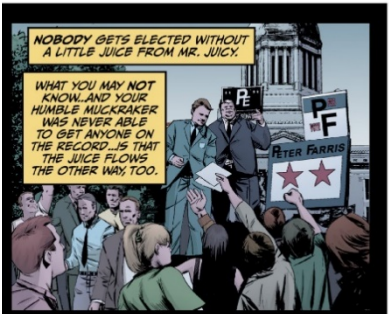
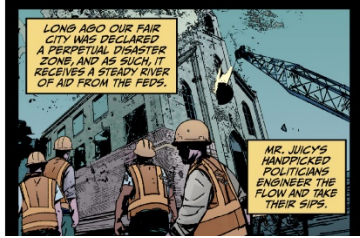
McCloud identifies six types of text-image combination, of which one is termed ‘text-specific’, in that the visuals rely heavily on verbal-based explanations (McCloud, 2006, p. 130). This is the type that best fits Panels 1-4, as, in these panels, the visuals merely serve to complement the detailed verbal texts. In Panel 2 the verbal text reads ‘NOBODY GETS ELECTED WITHOUT A LITTLE JUICE FROM MR. JUICY’, while the visual provides additional information about Berkshire, who is shown standing behind a politician as if in a supporting role. Something similar occurs in Panel 1 (‘KINGMAKER’, showing a social occasion), Panel 3 (‘DISASTER ZONE’, a shabby construction site) and Panel 4 (‘CHUG’, a man drinking). This complementarity between the two types of text indicates a synergistic collaboration between modalities, so that, in order to convert the visuals into audible verbal texts, the audiodescriber needs first to extract the key information from the accompanying verbal texts of the ST and then develop them as standalone descriptions.


The TT's strategy is to guide readers' attention from one panel to another by developing the existing monologues with movements that indicate a consistent perspective. Table 1 below presents a comparison of the ST and the TT to illustrate the additions made in the TT. It can be observed that the delivery of additional movements, description of the visuals and the verbal text from the ST takes a sequential pattern, with every description led by a movement of the same character, Leigh. Thus, readers follow Leigh's motions and focus of attention focus (as in Entry 4, ‘She held up a clipping’) and develop a comprehensive understanding of the contextual connection of all the visuals (e.g. AD Entries 2, 5, 8, 11), thereby positioning Leigh as a third-person reciter who guides readers through the key scenes without directly contributing to the background story of Berkshire's illegal profiting from political engagements. In this way, even though the visual descriptions are not tied to the verbal read-along as closely as juxtaposed images and texts are, they can regain coherence through their association with Leigh's narrative perspective.

Table 1. Comparison between Page 64 and its audio description

AD entries	Time Stamps	TT	Reference to ST
1	52:25-52:29	Leigh rifled through a pile of clippings, finding mentions of Berkshire in various places. Some of the clippings had pictures.	None

⁶ See Eisner (2008, p. 44) on how the shapes of boxes convey information about tense, text types, tones, and other elements.

2	52:30-52:38	Berkshire sitting at a trendy bar, red-hued martini at hand, surrounded by hangers-on and suits and cocktail dresses.	 <p>YOU KNOW MR. JUICY. HIS FINGERS ARE IN EVERY PIE IN THIS BURG, EVEN THE ONE YO' MAMA BAKED LAST NIGHT.</p> <p>MR. JUICY IS ALSO THE CITY'S UNOFFICIAL KINGMAKER.</p>
3	52:38-52:49	You know Mr. Juicy. His Fingers are in every pie in this burg. Even the one yo' mama baked last night. Mr. Juicy is also the city's unofficial kingmaker.	<p>YOU KNOW MR. JUICY. HIS FINGERS ARE IN EVERY PIE IN THIS BURG. EVEN THE ONE YO' MAMA BAKED LAST NIGHT. MR. JUICY IS ALSO THE CITY'S UNOFFICIAL KINGMAKER.</p>
4	52:49	She held up a clipping	None
5	52:50-53:08	that seems to be a victory Party for state representative, Peter Farris. There were huge campaign signs with the name Farris everywhere in the photo along with confetti and balloons. Peter Farris stood next to a podium as Berkshire looked on from the back, arms crossed and pleased with himself.	 <p>NOBODY GETS ELECTED WITHOUT A LITTLE JUICE FROM MR. JUICY.</p> <p>WHAT YOU MAY NOT KNOW, AND YOUR HUMBLE MUCKRAKER WAS NEVER ABLE TO GET ANYONE ON THE RECORD, IS THAT THE JUICE FLOWS THE OTHER WAY, TOO.</p> <p>PETER FARRIS</p>
6	53:08-53:21	Nobody gets elected without a little juicy from Mr. Juicy. What you may not know, and your humble muckraker was never able to get anyone on the record, is that the juice flows the other way, too.	<p>NOBODY GETS ELECTED WITHOUT A LITTLE JUCY FROM MR. JUICY. WHAT YOU MAY NOT KNOW, AND YOUR HUMBLE MUCKRAKER WAS NEVER ABLE TO GET ANYONE ON THE RECORD, IS THAT THE JUICE FLOWS THE OTHER WAY, TOO.</p>
7	53:21	Another clipping	None
8	53:21-53:31	Another clipping showed a ruined street in Arcadia as a wrecking ball smashed into a beautiful old church, sending bricks everywhere. Bored construction workers watched.	 <p>LONG AGO OUR FAIR CITY WAS DECLARED A PERPETUAL DISASTER ZONE, AND AS SUCH, IT RECEIVES A STEADY RIVER OF AID FROM THE FEDS.</p> <p>MR. JUICY'S HANDPICKED POLITICIANS ENGINEER THE FLOW AND TAKE THEIR SIPS.</p>
9	52:32-53:45	Long ago our fair city was declared a perpetual disaster zone, and as such, it receives a steady river of aid from the feds. Mr. Juicy's handpicked politicians engineer the flow and take their sips.	<p>LONG AGO OUR FAIR CITY WAS DECLARED A PERPETUAL DISASTER ZONE, AND AS SUCH, IT RECEIVES A STEADY RIVER OF AID FROM THE FEDS. MR. JUICY'S HANDPICKED POLITICIANS ENGINEER THE FLOW AND TAKE THEIR SIPS.</p>

10	53:46-53:49	Leigh picked up the first clipping and looks more closely at it now.	None
11	53:50-53:54	Berkshire and his hangers-on sipped martini in highballs that look as dark as blood.	
12	53:55-54:02	Mr. Juicy? He practically chugs it. Millions of dollars' worth. Allegedly, of course.	MR. JUICY? HE PRACTICALLY CHUGS IT. MILLIONS OF DOLLARS' WORTH. ALLEGEDLY , OF COURSE.

Similar examples can be found in the translation of the three juxtaposed mugshots on ST page 9, the flashbacks on page 80, and so on (Swierczynski, 2014), suggesting that this strategy is a frequent way of dealing with non-narrative panels when contextual implication in the ST makes it possible for the TT to present a clear one-character perspective. To establish connections between visually unrelated images, a narrative context is created by including additional descriptions that track a character's movements. As a result, the target audience perceives the TT through the filter of the character's perspective, rather than observing the panels from an objective viewpoint. To be more specific, by changing the perspective of the third-person narrator to present that of Leigh, as both the third-person monstator and reciter in audio, the different narration perspectives in the ST are unified under the same character in the TT. In this way, the AD can effectively establish connections between seemingly random visuals due to the consistent presence of the same character, even if there are sometimes discrepancies between the two versions. For example, Berkshire holds a glass of martini in AD Entry 2 but his hands are empty in the ST. The purpose of such in-panel adaptations would seem to be to add recurring elements for coherence (as mentioned above).

3.3 Multiple narrative panels: Additional scenes as cohesive devices

Apart from non-narrative transitions, the lack of strong panel-level semantic connections can also be found within narrative transitions. In contrast with the dynamic flow of audio description, narrative in comics has to be conveyed through a static series of images, requiring readers to fill in the gaps between panels (McCloud, 1994, pp. 62-68). Apart from location-setting panels (*orienters*) and time-prolongation panels (*prolongations*), a typical narrative arc in comics consists of four types of conjunctions: *establishers*, *initials*, *peaks* and *releases* (Cohn, 2013, pp. 70-77, see Table 2 below). Switching or omitting such key scenes can lead to a breach of the cohesive narration chain to various degrees (Cohn, 2012), although for artistic reasons (such as composition or pace control) such practices are considered legitimate in narrative comics. Another type of weak link between panels may

occur between one complete narrative arc and another (for example, two sequential causal results or double narrative threads, no matter whether they belong to the same major arc or not). In *X: Big Bad*, both types of weak transition are common, often used to accelerate a fight scene or other action scene with a small number of finely divided panels to indicate a very brief time duration and to form a contrast to larger close-ups. Additionally, as the entire issue consists of four chapters, each with multiple fight scenes and shifts in perspective, minor narrative arcs are consistently embedded within major arcs to enhance the complexity of the narration. Describing these transitions presents a challenge for AD.

The following discussion will present two examples for each type of weak panel transition, arguing that, when it is difficult to achieve cohesion by simply adding recurring items or contextualising with a consistent perspective, adding full transitioning scenes that include motions, settings, sound effects and even verbal lines is acceptable. When used appropriately, this strategy can enhance accessibility without disrupting the TT's narration.

Table 2. Narrative categories and conceptual structures (Cohn, 2013, p. 96)

Narrative Category	Conceptual Structure
Orienters	Superordinate context Location of event
Establishers	Introduction of referential relationship Passive state of being
Initials	Preparatory action Process Departing a Source of a path
Prolongations	Position on trajectory of a path Sustainment of a process Passive state (delaying)
Peaks	Culmination of event Termination of a process Interruption of event or process Reaching a Goal of a path
Releases	Wrap-up of narrative sequence Outcome of an event Reaction to an event Passive state of being

The first example is taken from pages 86 to 87 in the ST (Swierczynski, 2014) and the corresponding TT from 1:13:23 to 1:15:08 (Messner, 2021, Table 3 below). It demonstrates how the narrative arc, which was slightly fragmented in the ST due to the lack of transitions, is completed in the TT with additional elements. The two pages depict Leigh's attempt to escape from mobsters, creating an overall coherent narrative but with only a few key scenes missing. For clear reference, the panels in Figure 5 are marked with page numbers plus sequential numbers. P86-1 to 2 (*initial*: getting caught—*peak*: fighting back) and 3 to 5

(*establisher*: running into a crowd—*initial*: being targeted—*peak*: getting caught) form two minor narrative arcs with consistent cohesion respectively.





Figure 5. Pages 86-87 as an example for adding narrative structures within an arc (Swierczynski et al., 2014)

When placed within the larger arc of Leigh's escape from the scene, P86-1 to 2 together serve as an *establisher* for the following panels (3 to 5), as they form the starting-point of Leigh's escape by preparing background information about from where and whom she flees. However, when examining the transition between 2 to 3, there is a gap when Leigh makes her way out of the room with bookshelves and runs into a hall-like space with unpleasant-looking people surrounding her. In other words, there is a key scene missing from the narration that would support the *establisher*. In the comic version, the cohesion is compensated by the recurrence of the character Leigh and a consistent colour scheme (see Figure 5). The choice of similar deep greyish blue and green colours implies that the scenes represented in the two panels may not be taking place in the same room but certainly within the same building, while the character's running pose confirms that she is moving from one room to another but is yet to escape. While readers of the ST rely on these visual elements subconsciously to complete the narrative arc, the TT must compensate for this by utilizing auditory channels to provide the necessary information to its audience.

What the TT chooses to do is to add transitional scenes that function almost in the same way as the panels do. The descriptions of P86 2-3 and their matching STs are as follows:

Table 3. Comparison between page 86 and its audio description

AD entries	Time Stamps	TT	Reference to ST
1	1:13:42- 1:13:45	[s]he turned and swung the hard-cover book right into his face.	
2	1:13:45	*Female voice yelling*	None
3	1:13:46	*Male voice groaning*	GAK
4	1:13:46- 1:13:50	The edge of the book nailed Field Mouse right in the mouth. His mouth popped open, and blood flew.	
5	1:13:51- 1:13:53	Yeah. A real blockbuster.	YEAH. A REAL BLOCKBUSTER.
6	1:13:54- 1:13:57	*Female voice panting* Leigh dropped the book and spurred back into the dark hallway.	None
7	1:13:58- 1:14:01	She raced through a gauntlet of homeless who were curious about all the noise.	

By comparing the ST and the TT, we can see that AD Entry 6 serves the function of an establisher, as it not only provides transitioning moves, from Leigh hitting the mobster with a book to her running into a crowd, but also establishes the change in settings by specifically pointing out ‘the dark hallway’. As this addition focuses on tracing Leigh's movement rather than guiding the reader's attention to specific objects through her perspective, it can be considered an added scene rather than a perspective reinforcement. Additionally, the TT incorporates the sound effect of panting to make Leigh's actions audible.

This practice is similar to AD Entry 3, where the onomatopoeia in the ST is translated into non-verbal sounds to describe Field Mouse being hit. Thus, if back-translated, AD Entry 6 would require at least one transitional panel (possibly two panels for the two movements) that would clearly depict the consecutive movements and the change of scene, alongside possibly graphic onomatopoeic words. The translation of P86-5 to P87-1 also applies the same strategy, where the transition between ‘homeless people approaching Leigh’ and ‘the mobsters re-capturing Leigh’ is strengthened by ‘Field Mouse emerged from the bookstore behind her’ and ‘The other members of the Rat Squad ran up to Leigh and stopped’. Both examples illustrate the need for addition in this situation.

As for the transition between narrative arcs, the addition between pages 62 to 63 is another typical example of adding establishers. As seen in Figure 3, P62-1 to 5 is a very typical full narrative arc that contains a building setting (the *establisher*, 1),⁷ two men talking (the *initial*, 2), one man pointing at the other and chastising him (the *peak*, 3), and the other man leaving with X's shadowy figure staring from above (the *release*, 4) and *prolongation* (5). Directly moving out of this narrative arc, P63-1 switches to Leigh taking off her shoes without giving contextualisation as P62-1 does, which makes it an initial instead of an establisher. Even though P63-1 does provide hints that help make readers aware of the new setting (for example, furniture in a studio apartment and a warm, bright sky outside the window), these hints rely heavily on visual interpretations and must be organised for contextualisation. To fill in this gap, the TT prepares the target audience for the new arc by adding more movements of Leigh's arrival back home:

Door opening sound effect Leigh opened the front door of her shabby apartment. It was noon. She was dirty, banged up, and exhausted’ (Messner, 2021, 51:19-51:26).

These are all reasonable deductions, given that the next line in the script is ‘Leigh took off her shoes’ (ibid., 51:27-51:28). In other words, the TT gathers background information from the initial panel, and uses a movement to organise it into an establish panel-like description.

⁷ This is a very interesting panel for function categorization. Its visual depicts only the location, so it appears to be an orienter. However, since its dialogue boxes imply characters and their actions, such a referential relationship makes it an establisher instead of orienter.

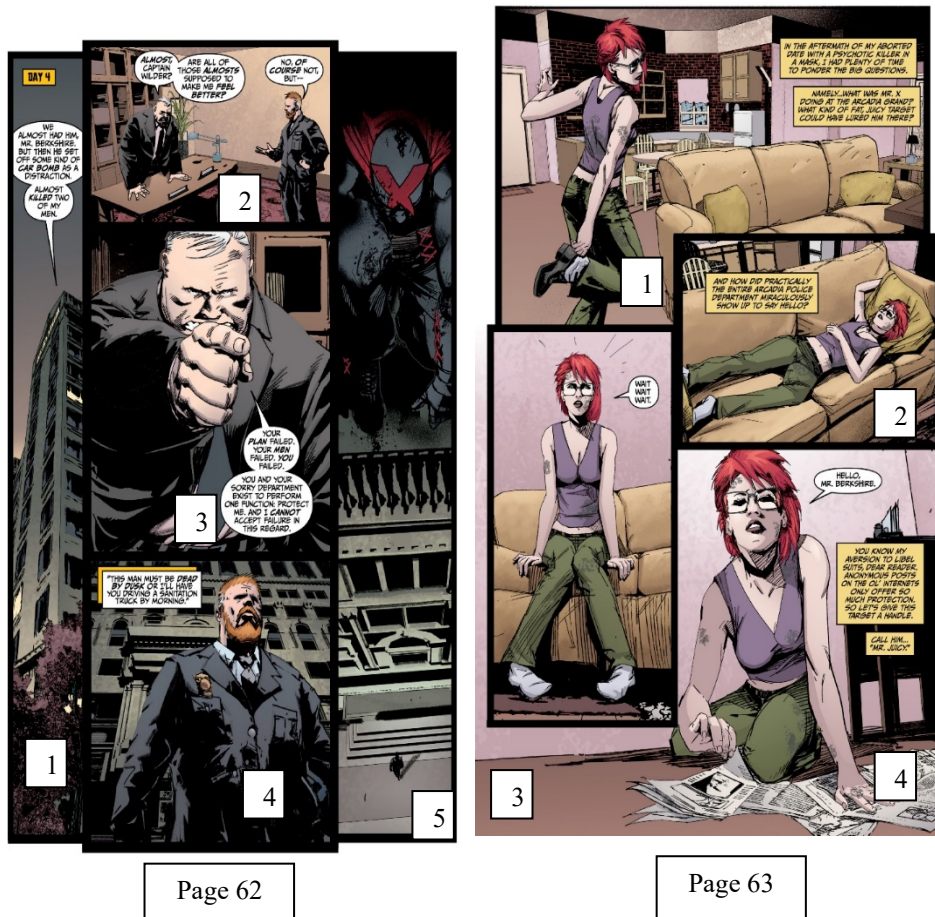


Figure 6. Pages 62-63 as an example for adding narrative

3. Conclusion

Through visual language analysis and a comparative case study of *X: Big Bad* (a commercial AD comic which adopts a 'movie-like' approach and achieves a highly coherent target text), this article has identified patterns in the translation methods used and concluded that certain norms exist that could contribute to comic AD translation in future. The findings suggest that, contrary to the assumption that AD comics must be utterly objective and that describers should refrain from interpretations, the TT's coherence is achieved precisely by mediating the discrepancy between the two types of multimodal texts, namely the verbal-visual ST and the audio-verbal TT.

Firstly, the TT emphasises important clues and recurring visual items with fixed, repetitive expressions and tends to add additional information for contextualization. When necessary, it also makes deductions about the existence of a certain item and repeats it, based on the contextual implications, even if it is not directly visible in the ST images. Secondly, for a few panels that are relevant to the main story arc but are not juxtaposed to form a narrative among themselves (a procedure which exists even in very typical narrative comics like this one), the TT turns to addition and creates a consistent character-as-narrator perspective that links the panels with the same context. Thirdly, for narrative panels that are weakly linked due to transitions between minor story arcs or for artistic reasons, the TT

fills in missing scenes with background introductions, character movements and sound effects. Combined, these added elements in the TT serve similar functions to the missing panel. By evaluating their contributions to the quality of the TT, it can be argued that even though such strategies could make the describers' mediation visible, the TT is made more accessible by providing a clearer logical chain; this would seem to make these interventions more acceptable.

This study is limited to a very defined translation goal and only examines the adaptations that are relevant to coherence. Thus, it intends merely to inspire solutions for dealing with certain challenges in comics AD, and in no way aims to fully respond to broader debates, such as whether the movie-like approach or the comic-like approach is generally a more acceptable choice. Notably, the adaptations in the TT are not limited to the categories mentioned in this essay, and the lack of translators' notes and readers' responses also limits its conclusions. Nonetheless, these limitations could provide a direction for future research into (for example) whether such norms develop through the influence of the ADs of more popular genres such as TV series or movies, or as a result of target audience responses to such marked interventions.

REFERENCES

- Brownie, B. and Danny Graydon (2016) *Bloomsbury Fashion Central - The Superhero Costume: Identity and disguise in fact and fiction* [online]. Available at: <http://www.bloomsburyfashioncentral.com/bfccoll/encyclopedia> (Accessed: 1 April 2023).
- Christopher, B. (2018) 'Rethinking Comics and Visuality, from the Audio Daredevil to Philipp Meyer's Life', *Disability Studies Quarterly*, 38(3) [online]. Available at: <https://doi.org/10.18061/dsq.v38i3.6477>.
- Cohn, N. (2013) *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London and New York: Bloomsbury.
- Eisner, W. (2008) *Comics and sequential art: principles and practices from the legendary cartoonist*. New York, London: W.W. Norton (The Will Eisner library).
- Gavaler, C. (2017) 'Refining the Comics Form', *European Comic Art*, 10(2), pp. 1–23 [online]. Available at: <https://doi.org/10.3167/eca.2017.100202>.
- GraphicAudio (no date) *Accessibility Statement* [online]. Available at: <https://www.graphicaudiointernational.net/accessibility-statement> (Accessed: 1 April 2023).
- Groensteen, T. and Miller, A. (2013) *Comics and Narration*. Jackson, United States: University Press of Mississippi [online]. Available at: <http://ebookcentral.proquest.com/lib/gla/detail.action?docID=1113447> (Accessed: 1 April 2023).
- Kingett, R. (2016) 'Finally, There's a Comic Book Store for the Blind', *Vice*, 9 August [online]. Available at: <https://www.vice.com/en/article/xdmd9d/comics-empower-comic-book-store-for-the-blind> (Accessed: 1 April 2023).
- Kleege, G. (2016) 'Audio Description Described: Current Standards, Future Innovations, Larger Implications', *Representations*, (135), pp. 89–101.

- Liu, C. – Repetition and Addition as Coherent Devices in Audio Description of Comics
Translation Matters, 5(1), 2023, pp. 102-120, DOI: https://doi.org/10.21747/21844585/tm5_1a6
- Kruger, J.-L. and Orero, P. (2010) 'Audio description, audio narration – a new era in AVT', *Perspectives*, 18(3), pp. 141–142 [online]. Available at: <https://doi.org/10.1080/0907676X.2010.487664>.
- McCloud, S. (1994) *Understanding comics: the invisible art*. 1st ed. New York: Harper Perennial.
- McCloud, S. (2006) *Making Comics: Storytelling Secrets of Comics, Manga and Graphic Novels*. New York: WmMorrowPB.
- Messner, E. (2021) *X: Big Bad* [MP3] [online]. Available at: <https://www.graphicaudiointernational.net/x-volume-1-big-bad.html>.
- Osolen, R.S. and Brochu, L. (2020) 'Creating an Authentic Experience: A Study in Comic Books, Accessibility, and the Visually Impaired Reader', *The International Journal of Information, Diversity, & Inclusion (IJIDI)*, 4(1) [online]. Available at: <https://doi.org/10.33137/ijidi.v4i1.32405>.
- Packer, J., Vizenor, K. and Miele, J.A. (2015) 'An Overview of Video Description: History, Benefits, and Guidelines.', *Journal of Visual Impairment & Blindness*, 109(2), pp. 83–94.
- Stainer, W. (2004) 'Pictorial Narrative', in M.-L. Ryan (ed.) *Narrative across Media: The Language of Storytelling*. Lincoln and London: the University of Nebraska Press.
- Swierczynski, D. et al. (2014) *X: Big Bad*. Milwaukee: Dark Horse [online]. Available at: <https://www.darkhorse.com/Books/22-521/X-Volume-1-Big-Bad-TPB>.
- Taylor, C. (2016) 'The multimodal approach in audiovisual translation', *Target. International Journal of Translation Studies*, 28(2), pp. 222–236 [online]. Available at: <https://doi.org/10.1075/target.28.2.04tay>.
- TPGi (2017) 'Accessible Comics!!! – Cordelia McGee-Tubb : #ID24 2017' [online]. Available at: <https://www.youtube.com/watch?v=zsQcOV0dZew> (Accessed: 1 April 2023).

About the author: Chunwei “Harold” Liu is a third year PhD student at the University of Glasgow, specializing in Modern Language and Culture, and supervised by Professor Laurence Grove and Dr. Enza de Francisci. After graduating in Translation Studies, as part of the Glasgow-Nankai University double-degree, Harold has developed a strong interest in the study of graphic literatures and inter-semiotic translation. Their PhD thesis focusses on Feministic and Queer studies in verbal-visual texts and popular cultures, especially in superhero narratives and their secondary writings and arts. Also, as a practitioner, they work in translation, interpretation, second-language education and digital illustration.