VERBAL-TO-VISUAL TRANSLATION BASED ON LINGUISTIC AND NARRATOLOGICAL MODELS: A POETRY-COMIC TRANSLATION OF *SAPPHO AND PHAON*

Chunwei Liu*
University of Glasgow

ABSTRACT**:** Poetry comic translation typically involves a verbal poem as the source text and a verbal-visual poetry comic as the target text. This innovative type of multi-modal translation is a typical example to study inter-semiotic conversion. Although scholarly attempts have been made to create linguistic-based verbal-visual translations, there is still a gap in discussing whether visual linguistic and narratological theories can be applied to build a practical translation model. To rationalise the translation from words to images, it is necessary to divide the process of comic translation based on standard comic writing processes. After looking for possible analytical linguistic models accordingly and critically incorporating them, the model I propose is mainly consisted of the academic achievements of Neil Cohn, Chris Gavaler, Will Eisner, Scott McCloud, Thierry Groensteen, J. A. Bateman and J. Wildfeuer, aiming to deal with the procedures in poem-comic translation such as text segmentation, layout design, narrative perspectives, and word-image conversion. Based on theoretical discussions and a translation practise of *Sappho and Phaon* (Mary Robinson, 1796), it is argued that an incorporation of current comic linguistic theories is feasible to overcome the challenges brought not solely by the discrepancies between verbal and visual language systems but also by the multi-modal nature of TT to a large extend.
KEYWORDS: Verbal-to-visual translation; Poetry comics; Visual language grammar; Visual narratology; *Sappho and Phaon*

1. Introduction

In a lecture given at the University of Edinburgh on March 5th, 2021, Paul Karasik, an experienced cartoonist, editor and educator, described and evaluated his practical experience of verbal-to-visual translation. The title of this lecture, "Gained in translation" suggests not only that Karasik has a clear awareness of the translational nature of his novel-based comic *City of Glass*, but also that the adaptation from verbal texts to visual or visual/verbal format has moved beyond the rarefied domain of the art and is now considered a valid subject for translation studies.

Poetry comics, an emerging artistic practice, is "an inextricable combination of both comics and poetry" (Chrissy Williams, cited in Daniel Elkin, 2016). In other words, the term refers to multimodal poetic creations that are either produced in or translated into visual-verbal language. In practice, artists have been experimenting with transforming poetry into the comic format since the 1960s, producing a wide range of poetry comics and even founding a non-profit translation project supported by Poetry Foundation, an independent literary organisation based in Chicago[1].

As a visual art form that involves transfer or conversion between verbal and visual elements, poetry comics have naturally begun to attract the attention of scholars working in the field of multimodality and translation. Derik Robertson, for example, has noted the

*2431662C@student.gla.ac.uk
[1] Poetry Foundation's official website: https://www.poetryfoundation.org/

similarity between the lines of a poem and the panels of a comic, arguing that both forms share an inherent segmentivity that offers great potential for the transfer of information between the two (2015, pp.1-2). He also differentiates between poetry comics and illustrated poems, on the grounds that the former requires a combination of words and images to transmit the necessary information (Robertson, 2015, pp. 2-4). This makes sense to anyone working with word-to-image conversion: the target text (TT) is aiming to be multimodal and so its language has to "force" the readers to interpret both verbal and visual elements together.

Another article, written by Brian Bates, further identifies poetry comics' unique contribution to specifying the meaning and chronological/spatial information in the source text (ST), by presenting a comparative study of two visual versions of *La Belle Dame san Merci* (2016). It illustrates how visual language can specify the abstract concept of time and space in the narration of poems. These two articles thus cover many of the aspects that a comic artist has to consider in the production of poetry comics, including scriptwriting, the representation of time and space, and ways of delivering information. However, when applied to actual translation practice, they do little to help the artist avoid ineffective arbitrary adaptations and make rational translation decisions. To fill in this gap, I thus propose the following research questions:

1. Can linguistic and narratological analysis be used to guide the translation of poetry comics?
2. Are the current theories sufficient to help the translator deliver certain features of the ST, such as rhythm and verb tense?
3. Can the mutual influence between verbal and visual representations be rationalised using visual linguistics or linguistic-related theories?

The issues that interest a translation studies scholar are not limited to analysing the final product. Rather, the entire span of the translation process is important, from the pre-translation (preparation) and in-translation (translation decisions) phases through to the post-translation (the reception or impact of a target text). However, the above-mentioned research questions are obviously much more closely related to the translator's preparation and strategies than to the reception and impact of translation products.

Considering that most poetry comic practitioners are not self-aware translators, I propose in this paper to take advantage of my own experience as a visual translator to extend the theoretical discussion to translation practice. That is to say, I aim to use one of my own experimental translation projects as research material for the construction of a multimodal translation model that can rationalise and orient translation from verbal poems into verbal-visual poetry comics. The work concerned is a visual translation of a sonnet cycle entitled *Sappho and Phaon*, which was specifically carried out for this purpose.[2] The translation of three of these sonnets is presented in the epigraph to this special issue.

[2] This ongoing research project is being carried out under the joint auspices of the Universities of Glasgow and Nankai.

Written by Mary Robinson and originally published in 1796, *Sappho and Phaon* is a series of romantic sonnets that depict the tragic love life of the ancient Greek poet Sappho (Craciun, A. and McLane, 2008). By choosing the first-person perspective, Robinson explores Sappho's sensitive yet passionate mind, discussing her psychological status before she committed suicide in despair at the unresponsiveness of Phaon, the man she loved (Robinson, 2008). This sonnet collection is arranged in a chronological order that traces a broad narrative line (falling in love—being forsaken—chasing love—failing—giving up on life). On the other hand, the content within each sonnet follows the protagonist's flow of thoughts and is thus non-narrative and even abstract in nature. Robinson intentionally uses ambiguous syntax and symbolic language to code Sappho's feelings, thoughts, and struggles, especially the almost philosophical struggle between sense and sensibility. In short, translating *Sappho and Phaon* into poetry comic format involves both narrative and non-narrative elements. The ST's complex poetic language poses various challenges to the translator, making it an appropriate material for poetry comic studies.

The first section of this article provides a method of segmenting the ST, taking into account the differences between narrative and non-narrative connections. In this process, the ST becomes a script that determines the number of panels and their contents based on the visual language grammar and the studies of cohesion in abstract comics, which will of course be subject to further revision in the next step. The second section, framing, is centred on deciding the layout design of each frame and how this can be used to recreate the ST's time duration and rhythm. Finally, the third stage makes the leap from words to images, deciding what elements of the ST can be visualised with semantic discourse representations. The whole translation procedure in this article is illustrated with extracts from *Sappho and Phaon*, and the argument is made that visual linguistic and narratological theories may be productively applied to regulate verbal-to-visual translation through processes of analysis and alignment.

2. Segmentation and transitions between panels

The first step of writing a comic or a manga, as agreed by most of the major drawing guidebooks, is to decide on a script that not only depicts the overall storyline but also assigns the contents to different panels (Hirohisa, 2013, p. 8; Lee and Buscema, 2010, pp. 116-120). Indeed, I believe that the primary task of poetry comic translation lies not with converting specific words into individual images but with finding an appropriate method of segmenting the ST. As Robertson (2015, pp. 1-2) notes, it is not sufficient to merely take the lines of a poem as natural panels, though he provides little guidance as to how it could be done.

For practitioners, the standard for ST segmentation tends to vary from one case to another. For instance, R. Kikuo Johnson uses the line breaks in the original poems as a rough guide, with one panel translating each line; Paul Hornschemeier emphasises the grammar of the ST rather than its layout, and Diane Wakoski segments the verses based on an interpretation of sense-groups (Johnson & Stallings, 2020; Hornschemeier and Ted Kooser,

2020; Wakoski, 2020)[3]. Thus, to theorise ST segmentation, it is necessary to identify a pattern that avoids both over- and under-segmentation.

In this section, the primary segmentation will follow the practice of traditional comics, which usually depict only a single movement in each panel (McCloud 1994, p. 110). Neil Cohn's theory of visual grammar (2013, 2018) will then be applied to identify movements. Based on contemporary construction grammar, visual language grammar uses traditional panels (and thus movements) as basic units of the visual language of the sequential arts, providing a possible inter-semiotic tool to translate between words and images at the panel level (Cohn, 2012, p.5; Cohn, 2018, p.7). Structured in accordance with the narrative functions of pragmatic linguistics, it assigns distinct roles to panels in order to fulfil functions such as scene-setting, initiation, prolongation, peak or culmination, and release (Cohn 2013, pp.413-452; 2018, pp.2-3). The syntax of visual language allows three types of inter-panel relationships, including canonical narrative schema (main clause), conjunction schema (conjunction), and head-modifier schema (adverbial modifier) (Cohn, 2018, pp.2-7). The schemas are arranged in a hierarchical tree structure, which allows complex verbal sentences (involving parenthesis, embedding, subordination, compounding etc) to be recreated visually (Figure 1).

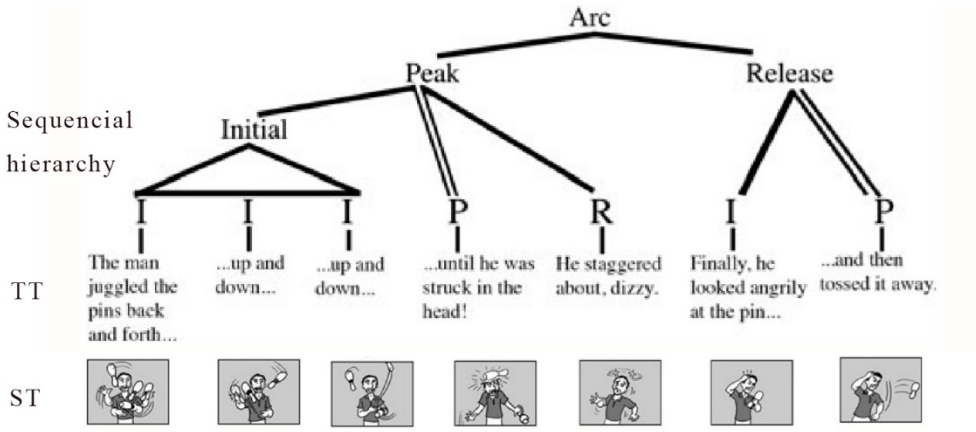

Figure 1 The conversion between visual and verbal language with visual language grammar

However, while Cohn's model is a feasible way of achieving alignment between verses and panels in narrative texts, the fact that visual language grammar is based on narratology suggests it is unsuitable for non-narrative texts or "illogical scrambled sequences", as he calls them (Cohn et al, 2012, p. 8). This means that it is not in itself enough to guide the translation process as it cannot describe the differences between sequential abstract comics proposed by Molotiu (i.e. the content of a comic is abstract, but its panels are arranged in a specific sequence), or the non-sequential "wallpaper effect" suggested by

[3] Johnson, Hornschemeier and Wakoski are poetry comic artists who have published their visual translations of verbal poems on *Poetry Foundation*. Please check the references for further details.

Thierry Groensteen (i.e. the panels in a comic are arranged together without an obvious sequence, creating an effect similar to a collage or wallpaper) (Molotiu, 2012, p. 89; Groensteen, 2013, p. 17). A recent study of abstract comics by Gavaler (2017, pp. 19-23) identifies six types of comics based on panel transitions and the visual elements within panels. These are:

1. *Representational narrative:* the most commonly accepted comic form in which the images are to some extent realistically portrayed and are linked together in order to narrate a plot.
2. *Abstract narrative:* abstract images that are juxtaposed in a certain order, forming a narration in the broader sense.
3. *Representational arrangement:* representational images that are arranged in an unordered but related manner.
4. *Abstract arrangement:* abstract images that are arranged in an unordered but related manner.
5. *Representational non sequitur:* juxtaposed representational images that are neither sequential nor logically related.
6. *Abstract non sequitur:* juxtaposed abstract images that are neither sequential nor logically related (Gavaler, 2017, pp. 19-23).

As the content of each panel is not the key priority in this section, Cohn's model can cover the first two types of narrative comics. On the other hand, the last four non-narrative categories achieve coherence by the juxtaposition of panels and repetition of visual elements. Thus, for non-narrative texts, segmentation should be placed between repetitions in the ST; the differences between narrative and non-narrative texts can also be identified with the categories of narratives and arrangements/non-sequitur.

By combining Gavaler's categories and Cohn's visual language grammar, the gap between a continual verbal text and segmented panels can be bridged with segmentation points noted as narrative + narrative (**NN**), narrative + non-narrative (**NNn**), non-narrative + narrative (**NnN**), non-narrative + non-narrative (**NnNn**) and embedded narrative or non-narrative language (**NN, NnN, NNn, NnNn**). In the following illustrations, segmented lines are regarded as sections and assigned into isometric squares rather than actual panels with artistic or rhetorical designs, which would require further supports to develop. **NN** is the most common transition between sections, and in Sonnet IV of *Sappho and Phaon*, lines 9-14 can serve as a typical example of a series of narrations (Figure 2):

9. Now, on a bank of Cypress let me rest; (**NN)**
10. Come, tuneful maids, ye pupils of my care, (**NN**)
11. Come, with your dulcet numbers soothe my breast; (**NN)**
12. And, as the soft vibrations float on air, (**NN)**
13. Let pity waft my spirit to the blest, (**NN)**

14. To mock the barb'rous triumphs of despair! (Robinson, 2008, p. 14)

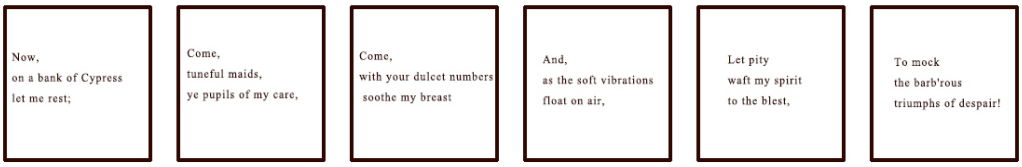

Figure 2 Segmentation of Sonnet IV Lines 9-14 into sections

In these lines, each line represents an action: “Sappho rests on a bank”, “maids come to Sappho”, “maids comfort Sappho with music and dance”, “music floats”, “pity passes Sappho’s spirit into the air”, and “mocking despair”. Sonnet IV lines 9-14 thus contains six narrative segments, which are placed in separate boxes and will become the foundation of the frame design.

NnNn combines two juxtaposed phrases or clauses that cannot form a story arc, while **NNn** and **NnN** refer to the situation where a non-narrative section is segmented from its narrative neighbours. These combinations are usually identified in fragmented poetic languages with ambiguous grammars and juxtaposed nouns or adjectives. For example, **NnNn** can be found within the last lines of Sonnets XX and XLII:

XX.14. Too proud to sue! (**NnNn)** too tender to resign!

XLII 14. "What suits with Sappho, **NnNn** Phoebus suits with thee!"

Even though the grammatical structures are not exactly parallel for each section, these two lines are segmented: the potential repetitions can be a tricky element to depict in a single panel but can serve as a mechanism to achieve coherence if separated into different panels. In Sonnet XX line 14, a repetition occurs in both non-narrative phrases, with “Proud” and “tender” being two static descriptions of Sappho’s’ thoughts. As love-related sentiments of the same person in a non-narrative roll, the two sections of XX 14 can be translated into different panels coherently, so at this stage, they are separated into two segments. As for Sonnet XLII line 14, this shares a similar situation where Phaon is seen to be equal to and suitable for both Sappho and Phoebus, which is also a descriptive comment, with the second-person reference to Phaon (“thee”) repeated. Thus, these two lines both belong to the **NnNn** category.

A typical example of **NNn** and **NnN** can be found in Sonnet IV, lines 4-9 (Figure 3):

4. And my chill'd breast in throbbing tumults rise? (**NNn)**
5. Mute, on the ground my Lyre neglected lies,
6. The Muse forgot, and lost the melting lay; (**NnN)**
7. My down-cast looks, my faultering lips betray,
8. That stung by hopeless passion, (**NN)**—Sappho dies! (**NN**)

9. Now, on a bank of Cypress let me rest; (Robinson, 2008, p. 14)

Figure 3 Segmentation of Sonnet IV Lines 4-9 as sections

Line 4 is segmented from line 5 as a divide between narration and non-narration. Line 4 contains a motion, "breast rising", while Line 5 can be categorised as a static description because it notices the status of Sappho's lyre as a variation of a "there-be" sentence. Line 6 lists two nouns: a forgotten Muse and a lost lay. As there is no direct imagery repetition in a goddess, a song, and a lyre, they are temporarily not separated into different sections. In lines 7 and 8, there is a narrative clause that contains only one movement "betray", so the segmentation is put between passion and the next movement "Sappho dies". To be more specific, "[m]y down-cast looks" is a parallel subject to "my faultering lips", and "[t]hat stung by hopeless passion" is the object of "betray".

These examples suggest that the first four categories of segmentation tend to keep integrated clauses (and sense-groups in fragmented sentences) in each panel regardless of the ST's layout. However, the fifth category – embedded narration/non-narrative phrases – may break the grammatical structure of the ST. In Sonnet XX, Lines 1-6 contain three embedded non-narrative sections (Figure 4):

1. Oh! I could toil for thee o'er burning plains; (**NNnN)**
2. Could smile at poverty's disastrous blow; (**NN)**
3. With thee, could wander 'midst a world of snow, (**NNnN)**
4. Where one long night o'er frozen Scythia reigns. (**NN)**
5. Sever'd from thee, my sick'ning soul disdains
6. The thrilling thought, the blissful dream to know, (**N^{N}N)**

Figure 4 Segmentation of Sonnet XX Lines 1-6 as sections

In this section, although the embedded narrative elements are not individual clauses or even meaning groups, they contain parallel repetitions and thus should be put in sub-sections. Apart from the narrations in lines 1, 3, and 5-6, namely "toiling", "wandering" and "disdaining", Phaon repetitively appears in adverbial phrases as "for thee", "with thee" and "sever'd from thee". On the one hand, as a parallel structure in lines 1 and 3, the repetitions

to the addressee indicate two possible non-narrative segments. In line 5, "sever'd" is a notional verb in its past participle form, suggesting a turn from non-narration to narration and should be separated from the former non-narrative section.

On the other hand, these three sub-sections can form an independent narrative arc independent of their main clauses. In Line 1, Sappho wishes she could walk over plains in pursuit of Phaon, and in the following lines, describes the psychological process of being with him before being soon separated from him. This is a typical narrative structure, as it can be described as initiation (perusing Phaon as in "for thee"), peak (being with Pahon as in "with thee"), and release (breaking up as in "sever'd from thee") (Figure 4). Thus, it is more adequate to treat them as sub-sections so they will not be overshadowed by the major sections, which can also imply that they happen in Sappho's spiritual world without confusing the reader by mixing them up with the main panels.

In stage 1, Gavaler's categories and Cohn's visual language grammar are thus combined to segment the ST and fit the results into sections and subsections. These sections will develop into the panels of the future TT. Narrative/non-narrative analysis is applied as a criterion for text segmentation, ensuring that each (ideally static) section only contains one element or one action to avoid unintentional over- or under-translation.

3. Framing and layout design

Unlike inter-panel transitions, layout design in comics has almost never been theorised in the language of the sequential arts, although it is commonly agreed that the topic should be covered by "visual language" in the broad sense (Xue, 2007, pp. 156-157). Textbooks for comic artists tend to describe the rules of layout design from a practical perspective, suggesting they should be "eye-catching" or "easier to edit" (Manhua Jifa Yanjiuhui, 2006, p. 46), and sequenced from left to right and top to bottom, with the first panel at the top left corner of each page (Idem). They also usually recommend that sizes and shapes of panels should vary to enhance the composition of frames and express emphasis (Xue, 2007, pp. 156-158, Manhua Jifa Yanjiuhui, 2006, pp. 46-47). Developed from practitioners' experience, these suggestions cannot provide a valid standard to define the characteristics of panels and thus are not specific enough to be regarded as secure guidelines for comic poetry translators. However, they do suggest that layout designs have the potential to re-create the ST's verbal features, as a unidirectional sequence of reading and rhetoric-like effects.

Will Eisner has studied the importance of comic frame design for expressing the concept of "time", arguing that "a comic becomes 'real' when time and timing is factored into creation" (2008, p. 26). Compared with traditional comics, modern comics allow multiple motions and a durable timespan to be expressed in a single panel, and the translator may merge sections from the first step in accordance with the need to present sequence, rhetoric and rhythms. Although comics are commonly expressed as static visual elements, readers can interpret motion and timespans from specific layout designs of panels and frames (Eisner, 2008, p. 28). For example, the comic artist may cancel the

gutters between multiple panels and enclose many static "slices" from a string of movements into one panel to "record a continued flow of experience" (Eisner, 2008, p. 39). McCloud (1994, pp. 110-115) identifies eight types of visual expressions that can capture time duration in comics: multiple panels in sequence, onomatopoeias, speech balloons, motion lines, multiple images, streaking effect, blurring effect, and subjective motions. The first category refers to the typical comic method in which a string of motions is visualised through multiple panels that contain a single action in each. The other seven are techniques for portraying multiple movements in the same panel, suggesting that the segments in the first stage are under-provisioned and can be revised and merged if necessary.

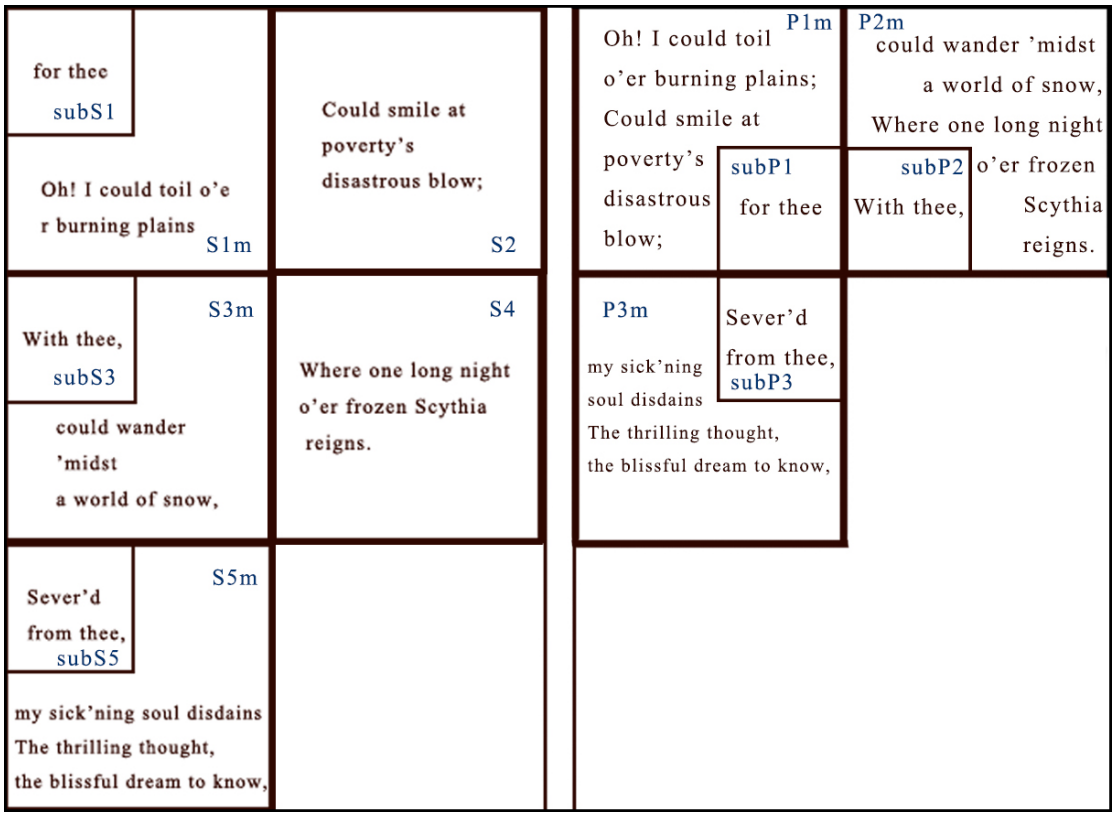

Figure 5 Comparison between the frame layouts of provisional segments and suggested panel design of Sonnet XX Lines 1-6

Eisner and McCloud's theories are particularly useful when the reading sequence is not in line with segmentation—especially with embedded sub-segments. As we have seen, the readers of comics are "trained" to read in a "left to right, top to bottom" direction (Eisner, 2008, p. 41); however, taking Sonnet XX lines 1-6 as an example, if a top-to-bottom sequence is applied, the four sub-sections will be too distant from one another (Figure 5), which could pose challenges for readers' comprehension. Closure (a cognitive effect that facilitates the understanding that a string of panels is a cohesive whole instead of individual images) only works when images are closely juxtaposed to each other (McCloud, 1994, pp. 62-68). Thus, it is likely that the sub-segments in this TT may not perceived as a cohesive story arc and will therefore have to be juxtaposed in the revision stage. In order to do this,

the major provisional segments given in Figure 5 (S1m and S2) can be merged as a revised panel (P1m), and provisional segments S3m and S4 become panel P2m. This revision can narrow down the space between sub-sections Sub1 and Sub3, which thus achieves cohesion between these imbedded sub-sections. In this case, the sub-sections are brought together to form a narrative structure and the major segments remain to follow the conventional sequence flow, so this layout design is more understandable than the original one. To translate the rhythmic language in the ST, frame structures can be adjusted to create a sense of rhythm; indeed, current studies suggest that both narrative and non-narrative comics can deliver a sense of rhythm with layout designs. In narrative comics, rhythm is usually presented as a variation in the time duration of motions and plots; while in non-narrative comics, rhythm is used more to create a poetic or musical effect (Groensteen, 2013, pp. 134-135).

On the one hand, panel shapes can influence time and rhythm (Eisner, 2008, p. 29). For narrative comics, Eisner presents various types of panel shapes which are commonly used to indicate tense (although these types are merely popular practice rather than established norms). For example, a straight-line box is usually interpreted as referring to the present tense, while a wavy box represents the past (idem). As the rhythm of narrative comics is built on the pace of time, panel shapes can be manipulated to influence the time flow and thus the rhythm of the TT. For narrative and non-narrative comics alike, a wavy, bubble-like panel usually indicates thoughts, while a star-shaped box suggests elements related to sounds or emotions (Eisner, 2008, p. 44). Non-standard panel borders and shapes might also be able to influence the rhythm directly. For example, the absence of borders can deliver a sense of time stretching on infinitely, and borderlines can create "a sense of heightened significance within the narrative structure" (Eisner, 2008, p. 45-49). Thus, even though Eisner's corpus cannot define quantitative qualities of panel shapes, his theory can provide a range of panel shapes for the translator to choose from in order to visualise pace, a timespan, or other features related to rhythm in the ST.

On the other hand, panels and multi-frames collectively function as the comics' beat or pulse. According to Groensteen (2013, p. 136), readers process a frame as both a totality and a collection, encompassing a group of panels. Thus, readers notice rhythms by comparing the spatial relationships between panels both as individual frames and as collections of panels (Idem). The more panels one frame contains, the higher the density of the whole page and thus the faster is the beat, and vice versa (Idem). Groensteen's argument illustrates that, to imitate the rhythm of the ST, the translator can alter the layout designs by selectively combining several segments to achieve a slow relaxed pace, which would make the rest of the TT comparatively faster and tenser.

For the translation of sonnets, which have a special rhythmic pattern, it is particularly important to manipulate the rhythm of the visual TT accordingly. A typical sonnet is composed of two parts: a thought or problem is proposed in the first eight lines and, after a turning point called *volta* (not a line but a structural feature), the issue is either addressed or challenged in the last six lines of resolution (Hurley and O'Neill, 2012, p. 76). As the first

section contains more lines, the first half of the TT will have more panels, which suggests a faster rhythm than the second section; this is in keeping with the practical experience that it is preferable to arrange smaller and tighter panels at the beginning of each frame (Groensteen, 2013, p. 136; Manhua Jifa Yanjiuhui, 2006, pp. 46-47).

For example, the volta in Sonnet XLII between lines 4-8 and lines 9-12 can be visualised with a change in panel size to mark a changing rhythm. The revised segments and the final panel design are illustrated in the form of comic strips below (Figure 6). The segments form a uniform pace from the top to S7-8, leaving a disproportionately large space for lines 13-

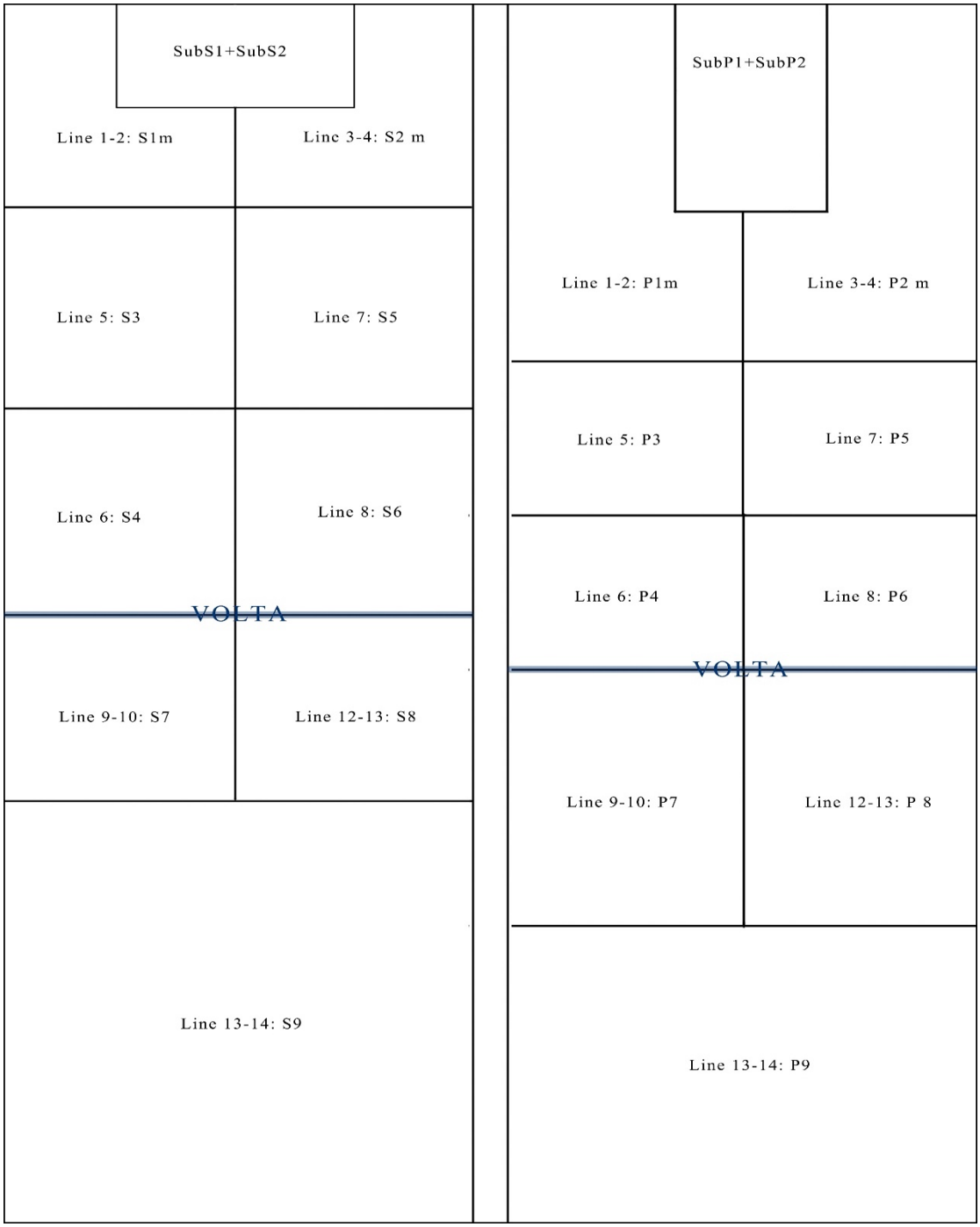

Figure 6 Comparison between the frame layouts of revised segments and suggested panel design of Sonnet XLII

14, which is not in keeping with the rhythm of the ST. What is more, the segments cannot mark the volta unless we introduce other visual elements such as thicker borderlines, which would divide the poetry comic into two frames instead of an integrated one. Thus, the size of S3-6 is cut down in an attempt to achieve a progressively slower pace with growing panel sizes, and the second part uses relatively larger panel sizes, marking the ST's volta and imitating the emphasis in ST with a change in the density of panels (Figure 6).

In Stage 2, the TT's layout design is used as a tool to transmit three of the ST's verbal language-based features: motion, time and rhythm. Theoretical segmentations from the first stage should be revised to adjust the arrangement and density of panels, achieving a durable timespan for the narrative coherence of subsections and the turn of volta. To conclude, the TT can imitate the ST's rhythm by adjusting panel sizes.

4.Word-image conversion

Just like verbal narrative texts, narrative comics also use narration perspectives to present storylines. According to Groensteen (2013, pp. 79-86), narrative devices in comics can be categorised into three types:

1. *Narrator*: the entity that controls the story.
2. *Reciter*: the character that provides the actual perspective to observe the story, who may be:
   - "In the background" (presenting the story from a silent point of view) vs "interventionist" (with specified narrations).
   - "Neutral" (i.e. unbiased) vs "involved" (implicated in the story).
   - "Reliable" vs "unreliable" in his/her presentation of the facts of the story.
3. *Monstrator*: the graphic character who performs the function of narration

This taxonomy of roles can help the translator to analyse what perspective the TT should apply to be in line with the ST, which is relevant not only to the overall story arc in the TT but to what items are depicted and how. For example, the whole collection of *Sappho and Phaon* is written from Sappho's first-person perspective, so overall the ST has a clear stance that marks Sappho as an autobiographic reciter, in Groensteen's terms. An autobiographic reciter, or I-as-character, is a self-representation of the first-perspective narrator that carries the narrative function (Groensteen, 2013, p. 98). In the ST, Sappho is a visible, interventionist and involved reciter, who directly refers to herself as "I" or "Sappho" and is activel implicated in the narration. To recreate this pattern, the TT not only draws on her first-person perspective but can also adopt the enunciation of a reciter. With narrative enunciation, the reciter's presence can be foregrounded and may be present in both the visual and the verbal text (Groensteen, 2013, p. 102). Thus, in this translation project, the TT is narrated from the first-person perspective, with Sappho's presence as both a narrator and a character in the verbal and visual modes.

The last step of the third stage is the conversion from words to images. To theorise the actual process of transferring verbal information into visual representations, this translation model makes use of Bateman and Wildfeuer's 2014 model of semantic

discourse representations. Semantic discourse representation theory is a verbal-linguistic system that examines how the "logic of discourse" is constructed by analysing semantic representations of meaning-making elements (Asher and Lascarides, 2003). Based on this theory, Bateman and Wildfeufer introduce visual representations into Asher and Lascarides' box-shaped model as elements that function as semantic representations in the same way as verbal texts (Figure 7). In this example, the first line of the box is the interpretation of the image, and the following two sections are graphic elements and action lines that contribute to the meaning construction (Bateman & Wildfeuer, 2014, pp. 189-193). The last section evaluates which representations would influence the meaning construction. By analysing verbal and visual elements as influential factors, this model can decrease the influence of subjective interpretation and narrow the gap between the vocabularies of the two semiotic systems, facilitating the conversion from words to images and helping to check the correspondence between verbal and visual texts in the TT.

Representation analysis works very effectively for concrete nouns and comparatively simple grammatical structures. For example, Line 5 of Sonnet XLII contains the information listed below (Figure 8). Line 5 is a description of a state of a heart ("warm") and is therefore

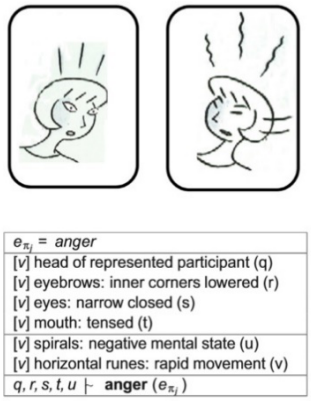

Figure 7 Discourse representation analysis of two consecutive panels
(Bateman, and Wildfeuer, pp. 191-192)

considered to be a non-narrative section. There are three major sense groups: "passions' flame", "heart" and "warm". The first two nouns are the two entities noted as x and y in Figure 8, and "warm" is the effect that x has on y. Thus, in the meaning construction of the ST, all three factors have contributed to the equation, "Y=warm", which represents the discourse meaning of Line 5. X and Y are both concrete nouns, so it is easy to convert them into pictorial vocabularies by referring to real-life objects.

In the analytical box of the TT' meaning, the overall connotation delivered in the TT is referred as Kπi' in Figure 8. Consisting of Kπi', factors X' and Y', which refer to the textual elements that contribute to a connotation, are combined and narrow down the range of possible effects—it should be a state of objects that is related to flame and can be applied to a heart, but it is not enough to deduce the exact connotation this visual expression

represents(the verbal text's connotation is referred in Figure 8 as eπi'). Its effect can be either positive, such as warm, or negative as burning, so verbal texts are necessary for the construction of meanings in this case. After adding the verbal translations of "passion's flame", "heart" and "warm" as meaning-making factors X', Y', and the meaning of visual elements as e πi' into the TT, the combination of verbal and visual factors in the multimodal TT are listed in the TT's connotation box (Kπi'). In this version, the effect of the TT is explicitly presented by the verbal expression "[warm]". Combined together, the visual elements, including the icons of a heart and flames (X', Y') and the three verbal phrases compensate each other and deliver the meaning of "passion's fire warms the heart", which conforms to the semantic meaning of the ST.

In certain situations, this model can even facilitate the inter-semiotic and inter-lingual translation of culturally specific items. For example, Phoebus, the god of the sun, in Sonnet XLII, Line 13 is translated as a stylised sun with a written name tag, so the two factors that influence the connotation of Phoebus (name, religious representation) are combined as a whole to fully deliver the ST without omission, expansion or footnotes.

In the final stage, the verbal contents assigned to each panel are converted into the multimodal TT. It illustrates that Groensteen's research on comic narratology can thus guide the analysis of the perspective of narrating and drawing the TT, while Bateman & Wildfeuer's model for semantic discourse representation can analyse the influential factors in the meaning construction process of both verbal ST and multimodal TT, bridging the gap between the two semiotic systems.

5. Conclusion

I have argued, based on my project of translating *Sappho and Phaon* into poetry comic form, that insights from visual linguistics and narratology can be applied to build a translational model that facilitates verbal-to-visual translation, to a large extent overcoming the challenges brought about by the discrepancies between verbal and visual language systems and the multimodal nature of the TT. The overall theoretical framework is a synthesis of the theories put forward by McCloud, Eisner, Groensteen and Gavaler to describe the units, characteristics and sequencing of the visual units, and two visual linguistic models developed by Cohn and Bateman & Wildfeuer respectively. These theories, originally designed to validate the analysis of images with verbal language, are applied here to first analyse the verbal ST and then convert it into a multimodal text. Thus, it can be argued that the alignment of the verbal ST and the verbal-visual TT in poetry comic translation can be achieved using visual linguistic models and categories borrowed from visual narratology, valid for both narrative and non-narrative aspects of the ST.

Given the potential of this translation model for use with different semiotic modes, I hope that it may inspire further multimodal translation studies. In future research, more genres of verbal texts and graphic literature could be involved in order to broaden the range of research materials. It would also be interesting to test whether this framework can function in the opposite direction: translating from visual to verbal.

REFERENCES

Primary Sources

Elkin, D. (2016) *Parsing poetry comics - an interview with Chrissy Williams*, *Comics Bulletin*. Available at: http://comicsbulletin.com/parsing-poetry-comics-interview-chrissy-williams/ (Accessed: 15 August 2021).

Hirohisa T. (2013) *マンガの学校 <1> 短編マンガの描き方 [Manga School <1> Techniques for Short Comics]*. Translated by Ding Lian. Beijing: China Youth Press.

Hornschemeier, P. and T. Kooser (2020) *The poem as comic strip #5 by Paul Hornschemeier, Ted Kooser*, *Poetry Foundation*. Available at: https://www.poetryfoundation.org/articles/68935/the-poem-as-comic-strip-5 (Accessed: 25 February 2020).

Johnson, R. K. and A. E. Stallings (2020) *The poem as comic strip #6 by R. Kikuo Johnson, A. E. Stallings*, *Poetry Foundation*. Available at: https://www.poetryfoundation.org/articles/69016/the-poem-as-comic-strip-6 (Accessed: 24 February 2020).

Lee, S. and J. Buscema (2010) *How to draw comics the Marvel way*. Translated by Meng Xianrui. Beijing: Posts & Telecom Press.

Manhua Jifa Yanjiuhui [Association of Comic Techniques] (2006) *Katong Manhua Huihua Jifa [Techniques for Drawing Comics and Mange]*. Hangzhou: China Academy of Art Press.

Robinson, M. C. (2008) *Sappho and Phaon: in a series of legitimate sonnets*. London: Forgotten Books.

Wakoski, D. (2020) *Belly dancer by Diane Wakoski*, *Poetry Foundation*. Available at: https://www.poetryfoundation.org/poems/46659/belly-dancer (Accessed: 25 February 2020).

Xue L. (2007) *Shaonv Manhua: Katong Manhua Rumen [Chick Comics: A Beginner's Guide to Manga and Comics]*. Wuhan: Changjiang Literature & Art Press.

Secondary Sources

Asher, N. and Alex Lascarides (2003) *Logics of conversation*. Cambridge: Cambridge University Press.

Bateman, J. A. and Wildfeuer, J. (2014) 'A multimodal discourse theory of visual narrative', *Journal of Pragmatics*, 74, pp. 180–208. doi: 10.1016/j.pragma.2014.10.001.

Bates, B. (2016) 'Graphic and digital keats: "la belle dame sans merci"' in poetry comics', *Reconstruction: Studies in Contemporary Culture*, 16(1). Available at: http://link.galegroup.com/apps/doc/A484096645/EAIM?u=glasuni&sid=EAIM&xid=eb9f7624 (Accessed: 10 June 2019).

Craciun, A. and McLane, M. N. (2008) 'Romantic poetry, sexuality, gender', in McLane, M. N., *The Cambridge Companion to British Romantic Poetry*. Edited by J. Chandler. Cambridge: Cambridge University Press, pp. 155–177. doi: 10.1017/CCOL9780521862356.009.

Cohn, N. *et al.* (2012) '(Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension', *Cognitive Psychology*, 65(1), pp. 1–38. doi: 10.1016/j.cogpsych.2012.01.003.

Cohn, N. (2013) 'Visual narrative structure', *Cognitive Science*, 37(3), pp. 413–452. doi: 10.1111/cogs.12016.

Cohn, N. (2018) 'In defense of a "grammar" in the visual language of comics', *Journal of Pragmatics*, 127, pp. 1–19. doi: 10.1016/j.pragma.2018.01.002.

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. doi: 10.3167/eca.2017.100202.

Groensteen, T. (2013) *Comics and narration*. Translated by A. Miller. Jackson: The University Press of Mississippi (Formes sémiotiques).

Hurley, M. D. and O Neill, M. (2012) *The Cambridge introduction to poetic form*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511982224.

Karasik, P. (2021) "Gained in translation: adapting literary texts into comics", University of Edinburgh, 21 April 2021.

McCloud, S. (1994) Understanding comics: the invisible art. 1st, Harper Perennial, ed. New York, N.Y: Harper Perennial.

Molotiu, A. (2012) "Abstract form", *Critical Approaches to Comics: Theories and Methods*. pp. 88–100, Oxford: Routledge.

Robertson, D. (2015) 'Justification of poetry comics: a multimodal theory of an improbable genre', *The Comics Grid: Journal of Comics Scholarship*, 5(1). doi: 10.5334/cg.bg.

About the author: Chunwei "Harold" Liu is a first year PhD student at the University of Glasgow, specializing in Modern Language and Culture. After graduating in Translation Studies, as part of the Glasgow-Nankai University double-degree, Harold developed a strong interest in the study of graphic literatures and inter-semiotic translation. Their PhD thesis focuses on Feministic and Queer Studies in verbal-visual texts and popular cultures, especially in superhero narratives and their secondary writings and arts. As a practitioner, they work in translation, interpretation, second-language education and digital illustration.