

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

A multimodal social semiotic approach to the analysis of manga:

A metalanguage for sequential visual narratives

Cheng-Wen Huang

HNGCHE004

A dissertation submitted in fulfillment of the requirements for the award of the degree of
Master of English

Faculty of the Humanities

University of Cape Town

2009

COMPULSORY DECLARATION

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature: _____ Date: _____

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr Arlene Archer, whose patience, kindness and academic experience have been invaluable to me. I am indebted to her critical assistance, thoughtful feedback and encouragement throughout the research period. Her financial support is also greatly appreciated.

I would like to extend my thanks members of the Multimodality in Education research group, Dr Arlene Archer, Dr Marion Walton, Dr Rachel Weiss, Terri Grant, Franci Cronje, Medee Rall, Nicola Pallitt, for their insight and useful comments on the research. I am especially grateful to Mariam Essack for editing the dissertation and Shabnam Parker for her useful feedback. In addition, I wish to thank my brother, Hsin-Chi Huang for introducing me to manga and helping me to find data for the research. Hsin-Hung Huang's assistance in compiling the dissertation is also hereby acknowledged. Special thanks are extended to my parents, Hsiu-Lan and Teng-Yuan Huang, who have been a constant source of support during my academic career.

Abstract

This study contributes towards an understanding of the nature of sequential visual narratives, how different semiotic resources may be employed to construct a visual narrative and how sequence of images may be developed. Over the years, extensive research has been undertaken in the area of still images. However, the particularities of meanings made in sequential images remain relatively unexplored. The significance of the study is that it contributes towards an understanding of sequential narratives by proposing a metalanguage for manga.

The term 'manga' refers to comics that originate from Japan and it is currently a trend in popular culture worldwide. Certain conventions employed in manga are different from that of Western comics. Using the proposed metalanguage, this study identifies the representational resources used in manga and examines how they are used to construct a visual narrative. The metalanguage is grounded in Kress and van Leeuwen's (1996) work produced in *Reading Images: The Grammar of Visual Design* and Matthiessen's (2007) concept of rhetorical relations in images.

The theory that underlines the study is multimodal social semiotics which assumes that texts are composed of a combination of representational resources. These resources are always socially situated, produced in a particular cultural, social and historical context. The theory supports the view of comics as a genre and makes it possible to attribute the differences between manga and Western comics to the social and cultural practices of the East and West. This study challenges the tendency in narrative tradition to favour verbal narratives over non-

verbal narratives by demonstrating that different representational resources employed in manga have distinct narrative functions and that they contribute to the meaning of the narrative in different ways. Moreover, meaning is derived from an integration of all the representational resource.

The study concludes by looking at the implications of using the metalanguage in interrogating other visual narratives. The New London Group's (2000) notion of 'designs of meaning' proposes that representational resources are like design resources. Individuals employ these resources in particular ways to produce particular texts. A social theory of genre highlights the overlapping nature of genres. Drawing on these concepts, this study argues that a metalanguage which can discuss different forms of meaning can also assist individuals to see the similarities between genres by foregrounding the use of conventions. From this perspective, it is possible to use the metalanguage to interrogate other visual narratives such as storyboarding.

Contents

Acknowledgements	i
Abstract	ii
List of Figures	vi
List of Tables	vii
Chapter One: Introduction.....	1
1.1 Background	1
1.2 Aim and research questions	3
1.3 Rationale	4
1.3.1 Rationale for a metalanguage for visual narratives	4
1.3.2 Rationale for using manga as the text of analysis	6
1.4 Overview of the thesis	10
Chapter Two: Theoretical Framework.....	12
2.1 Overview of chapter	12
2.2 Multimodality and the communication landscape	12
2.3 Multimodal social semiotic theory	15
2.4 A multimodal social semiotic approach to genre	19
2.5 Characteristics of an accessible metalanguage	23
2.6 Genres as ‘designs’	26
2.7 Comics from a social semiotic perspective	27
2.8 A social theory of genre	34
2.9 Narrative as a genre	39
2.10 Final comments	43
Chapter Three: Methodology.....	45
3.1 Overview of chapter	45
3.2 Overview of research method	45
3.3 Data	46
3.4 Framing the data: Labov’s narrative structure	48
3.5 Method of analysis	49
3.5.1 A metalanguage for manga	50

Chapter Four: <i>Naruto</i> from a social semiotic perspective.....	70
4.1 Introduction	70
4.2 Abstract	71
4.3 Orientation	76
4.4 Complicating Action	93
4.5 Evaluation	106
4.6 Resolution	113
4.7 Final comments	119
Chapter Five: The implications of the study.....	120
5.1 Overview of chapter	120
5.2 Semiotic resources and their affordances	120
5.3 Mixing Logics	127
5.4 The influence of social and cultural practices on manga conventions	129
5.5 The possible implications of using a metalanguage of manga in interrogating other visual narratives	133
5.6 Using a metalanguage of manga to examine storyboarding	134
5.7 Final comments	137

List of figures

- Figure 1: Representing a social experience (Kishimoto 2007: 86)
- Figure 2: Portraying social experiences through body posture, gesture and facial expression (Kishimoto 2007: 19)
- Figure 3: Extension through tilting an image (Kishimoto 2007: 24)
- Figure 4: A shift in time and transition through transition (Kishimoto 2007: 18)
- Figure 5: Projection through a voiceover (Kishimoto 2007: 17)
- Figure 6: A first person subjective point of view (Kishimoto 2007: 58)
- Figure 7: A third person point of view (Kishimoto 2007: 16)
- Figure 8: An omniscient point of view (Kishimoto 2007: 18)
- Figure 9: The index finger demands interaction from the viewer (Kishimoto 2007: 12)
- Figure 10: A canted angle signifies a world that is distorted (Kishimoto 2007: 35)
- Figure 11: The diagonal frames simulates the action (Kishimoto 2007: 31)
- Figure 12: Jagged speech frames signify the intensity and volume of voice (Kishimoto 2007: 14).
- Figure 13: “Homospaciality” (Kishimoto 2007: 50)
- Figure 14: Typography as an important resource to conjure up the sound effect (Kishimoto 1999: 9, 14, 45)
- Figure 15: English translations of the Japanese sound effects (Kishimoto 2007: 9, 14, 45)
- Figure 16: The original *Naruto* and the fan English translated version (Kishimoto 1999: 4)
- Figure 17: Abstract from the English edition (Kishimoto 2007: 4)
- Figure 18: Wide shot establishing the location (Kishimoto 2007: 9)
- Figure 19: Specifying an image through elaboration (Kishimoto 2007: 9)
- Figure 20: Orientation, establishing the setting (Kishimoto 2007: 9)
- Figure 21: Orientation, establishing the characters and their situation (Kishimoto 2007: 10-11)
- Figure 22: Close-up shot establishes a sense of intimacy (Kishimoto 2007:10)
- Figure 23: Re-establishing the setting through an omniscient point of view (Kishimoto 2007: 10)
- Figure 24: A medium shot (Kishimoto 2007: 10)
- Figure 25: The notion of ‘us’ and ‘them’ is established through the foreground/background continuum (Kishimoto 2007: 10)
- Figure 26: Naruto swoops into view (Kishimoto 2007: 11)
- Figure 27: Expanding an image through extension (Kishimoto 2007: 11)
- Figure 28: Expanding an image through projection (Kishimoto 2007: 11)
- Figure 29: A wide shot re-establishing the location (Kishimoto 2007: 29)
- Figure 30: The use of elaboration creates tension in the narrative (Kishimoto 2007: 29)
- Figure 31: A moment of comic relief (Kishimoto 2007: 29)
- Figure 32: Close-up shots depicting reaction after outburst (Kishimoto 2007: 29)
- Figure 33: Frames establish tempo (Kishimoto 2007: 29)
- Figure 34: Ominous mood established through framing (Kishimoto 2007: 30)
- Figure 35: A moment of revelation (Kishimoto 2007: 30)
- Figure 36: An omniscient point of view (Kishimoto 2007: 30)
- Figure 37: The cause of the complicating action revealed (Kishimoto 2007: 30-31)
- Figure 38: A split frame (Kishimoto 2007: 31)
- Figure 39: Expanding a sequence of images through extension (Kishimoto 2007: 31)
- Figure 40: Complicating action (Kishimoto 2007: 30-31)

- Figure 41: Evaluation (Kishimoto 2007: 48)
- Figure 42: Naruto overhears a conversation between Iruka and Mizuki
(*Kishimoto 2007: 48*)
- Figure 43: The white background signifies Naruto's isolation (*Kishimoto 2007: 48*)
- Figure 44: Speech that overlaps frames (Kishimoto 2007: 48)
- Figure 45: Developing the sequence of images through projection (Kishimoto 2007: 49)
- Figure 46: Pacing the narrative through wordless panels (Kishimoto 2007: 49)
- Figure 47: Climax (Kishimoto 2007: 50)
- Figure 48: Actions prior to Resolution (Kishimoto 2007: 51-52)
- Figure 49: Resolution (Kishimoto 2007: 53)
- Figure 50: 'The art of the doppelganger' (Kishimoto 2007: 54-55)
- Figure 51: Illustrating excitement through colour (Kishimoto 2007: 30)
- Figure 52: Lines and shapes as semiotic resources (Kishimoto 2007: 29)
- Figure 53: The narrow frames suggest an ellipsis (Kishimoto 2007: 58)
- Figure 54: Elements of a storyboard (Tumminello 2005: 5)
- Figure 55: Transforming manga into a storyboard

List of tables

- Table 1: The representational metafunction
- Table 2: The interactive metafunction
- Table 3: The compositional metafunction

Chapter One: Introduction

The dominance of visual images in our communications landscape has brought about a surge of interest in the study of images. Visual narratives are a research area that requires further exploration. This study proposes a metalanguage for a particular visual narrative, manga. Manga is Japanese comic art and it is a trend in today's popular culture. It is chosen as the text of analysis because of its current popularity and because the conventions employed in manga are comparable to other visual narratives. This makes it an ideal text to build a metalanguage on as it points to the possibility of extending the metalanguage to account for other visual narratives.

The study intends to use the metalanguage for manga to analyse *Naruto*, a manga narrative. The aim is to investigate how representational resources can be employed for narrative purposes. By examining the narrative functions of various representational resources used in manga, the study seeks to challenge the logocentric bias in narrative tradition.

1.1 Background to research

Although the origin of manga is noted to date back to seventh century caricatures found in the Horyuji Buddhist temple (Rubinstein-Ávila and Schwartz 2006; Ito 2005), manga, in its modern form, is said to have emerged in the 1950s (Kinsella 2000). Manga narratives are usually categorised into four kinds: 'shonen' for boys, 'shojo' for girls, 'seinen' for adults, 'rediiisu komikku' for ladies. The data used in this study, *Naruto*, is a shonen narrative. Production costs for manga are low because, except for the covers, it is usually printed in

black and white. In Japan, manga's success can be attributed to its diverse genre and cheap production costs.

Manga spread to the West in the early 1980s accompanied by 'anime', Japanese animation. In the United States, manga sales were noted at an estimate of US\$100 million in 2003 (Rubinstein-Ávila and Schwartz 2006). This number rose to between US\$175 million to US\$200 million in 2006 (Wikipedia 2009). The art form has inspired a trend in manga-style comics in the West such as 'la nouvelle manga' in France and 'Amerimanga' in the United States. In the United Kingdom, a publishing house, *Self Made Hero*, has taken advantage of the hype around manga and produced a series of Shakespeare plays in manga-style. These include plays such as *Romeo and Juliet*, *Hamlet*, *Richard III* and *The Tempest*. *Self Made Hero* claims that the manga adaptations of Shakespeare is intended to make the plays "accessible...through a medium that's increasingly popular with kids" (Eason 2007).

In the past, certain educators have tended to view popular culture texts as lacking content and not inappropriate for academic use. Today, however, literacy researchers are validating these texts, arguing that popular culture texts can provide access for literacy development (Gee 2003; Rubinstein-Ávila and Schwartz 2006). This view is supported by theories such as multimodality, multiliteracies, cultural studies, theories that have emerged in the twentieth century as a result of technological, social and cultural changes in society (Cope and Kalantzis 2000; Kress 2003; Gee 2003).

I became interested in analysing manga after noticing its increasing popularity in South Africa, particularly among university students. Multimodal social semiotics is a theory which can account for social and cultural influences in texts as well as meanings made in

multimodal texts, thus the decision was made to analyse manga from a multimodal social semiotic perspective. The idea of constructing a metalanguage for manga and the possibilities of extending this metalanguage to account for other visual narratives followed after research indicated a need for a metalanguage for visual narratives.

1.2 Aim and research questions

The aim of the study is to design a metalanguage of analysis for manga and to explore the possible implications of using the metalanguage in interrogating other visual narratives. By visual narratives, this study refers to stories told by means of static images in sequence, for example, comics (strips or books), picture books and storyboards. A metalanguage is “an educationally accessible functional grammar” that describes various forms of meaning available for meaning-making (NLG 2000: 24). It is a language which allows for a critical analysis of semiotic systems. The study aims to devise a metalanguage that is adequate for a theoretical analysis of manga but at the same time accessible to students.

In proposing a metalanguage for manga, the study will investigate the nature of manga narratives, how various semiotic resources are employed to create meaning in manga. The following research questions are essential in guiding the study:

- 1) How are semiotic resources used to narrate a story in manga?
- 2) What are the possible implications of a metalanguage for manga in interrogating other visual narratives?

1.3 Rationale

This study aims to design a metalanguage of analysis for manga and to investigate the possible implications of using the metalanguage in interrogating other visual narrative genres. The significance of the study is that it contributes towards theorising a metalanguage for visual narratives and at the same time it explores the pedagogic efficacy of using popular genre in literacy education.

1.3.1 Rationale for a metalanguage for visual narratives

Metalanguages are sets of grammars that describe semiotic resources and how they function to construct meanings. Semiotic resources refer systems of semiotic forms that we use to make meanings (Baldry and Thibault 2006). In other words, they are the resources that we use to communicate. To discuss and analyse semiotic resources, a particular set of vocabulary or a metalanguage is necessary. Metalanguages are important because they make it possible to describe the functional aspects of a semiotic system (Unsworth 2007). In doing so, they enable one to understand how the resources within the system function to make meaning. Over the years, numerous theorists have proposed metalanguages to describe various semiotic resources. Halliday's systemic functional linguistics (SFL), for instance, is a metalanguage that describes the resources used to make meaning in language. The twentieth century in particular saw an increased interest in the study of semiotic resources other than language. Above all, the analysis of visual images has been extremely popular and various metalanguages have been proposed for various types of visual images. For example, Kress and van Leeuwen's (1996) 'grammar for visual design' has been most successful as a metalanguage for still, single-framed images. In addition, both Unsworth (2006) and

Martinec and Salway (2005) have proposed metalanguages to describe the meaning-making resources of language and image interaction. More recently, there have been attempts to describe the meaning-making resources of sequential images (Matthiessen 2007; Baldry and Thibault 2006; Lim 2007). However, a comprehensive framework for meanings constructed through sequential visual narratives has yet to be developed. This study seeks to unify and build on the works by the above theorists, in particular the works of Kress and van Leeuwen (1996) and Matthiessen (2007), and propose a framework and a metalanguage for visual narratives that is accessible to students.

In theorising a metalanguage for manga, this study will contribute towards theories of multimodality. 'Multimodality' is a theory of communication that "accounts for the increasing multiplicity of modes of meaning-making, and theorizes the links between shifting semiotic landscapes, globalization, re-localization, and identity formation" (Archer 2006a: 451). 'Modes' describe semiotic resources that are culturally and socially shaped for representation and communication, for example, language, image and gesture (Kress 2003). In social semiotics, all modes are seen as possessing particular meaning-making potentials.

As Kress writes,

semiotic modes have different potentials, so that they afford different kinds of possibilities of human expression and engagement with the world, and through this differential engagement with the world they facilitate differential possibilities of development: bodily, cognitively, affectively (2000: 157).

Western literary tradition, however, has largely ignored the meaning-making potentials of other modes and privileged the linguistic mode above any other mode. This has created deep-rooted beliefs that knowledge and meaning can only be expressed through language. Narrative theories, for example, have been so exclusively shaped and derived from theories grounded in language that according to Prince (1988) narrative analysis itself favours verbal

over nonverbal narratives. Similarly, Ryan notes “[i]t seems clear that of all semiotic codes language is the best suited to storytelling. Every narrative can be summarized in language, but very few can be retold through pictures exclusively” (2004: 10). According to Ryan, this is a result of the inability of images to make propositions and causal relations. ‘Proposition’ refers to the act of “picking a referent from a certain background and of attributing to it a property also selected from a horizon of possibilities” (Ryan 2004: 10). Ryan’s observation suggests that theories that are used to analyse narratives are mostly language-based. However, various researchers (Archer 2006b; Kress et al 2001) have noted that different modes offer different possibilities for representation. It cannot be appropriate to examine a semiotic mode using a theory intended for another semiotic mode. The tendency in narrative tradition to favour verbal over nonverbal narratives is a result of a lack of understanding of the potentials of other modes in meaning-making. In today’s multimedia environment, it is evident that narratives are not limited to the linguistic mode alone. Narratives are told across media, through moving pictures such as television and film; sequential images such as comic strips and picture books. Therefore, in order to account for contemporary narratives, it is necessary to look beyond language-based narratives. Multimodality is an appropriate approach to multimodal narratives because it recognises that different modes have different meaning-making potentials.

1.3.2 Rationale for using manga as the text of analysis

Manga is chosen as the text of analysis for three reasons. The first reason is due to manga’s status as a popular cultural text. In the last decade or so there have been changing perspectives on popular culture and literacy as the pedagogic efficacy of using popular cultural texts in schooling becomes more apparent to educators (Alvermann and Heron 2001;

Norton and Vanderheyden 2004; Gee 2003; Rubinstein-Ávila and Schwartz 2006). Norton and Vanderheyden's (2004) research with second language learners and *Archie* comics, for example, demonstrates how popular cultural texts can function to cross cultural and linguistic barriers and allow both second language and first language speakers to engage in active class discussions. Popular cultural texts are reflective of ideologies and meaning-making mechanisms of our current society. According to Foucault, "each society has its regime of truth, its 'general politics' of truth: that is, the types of discourse which it accepts and makes function as true" (1995: 131). 'Discourse' describes "a way of signifying a particular domain of social practice from a particular perspective" (Fairclough 1995: 14). This suggests that truth is not a given, but a social construct that exists in time and space. In other words, truth is based in the society, culture and historical period of its creation. Students who are raised in the current era may find it increasingly difficult to connect with texts of the past because they reflect a different social truth to the one we currently occupy. Since popular cultural texts reflect a world in which students are immersed, they can serve as valuable teaching resources.

Furthermore, in the current era of change, it is no longer reasonable and sufficient to only privilege literary genres (Kress 2003). Texts are always sites of struggle. The struggle for representation and which texts will 'count' in academic literacy practices may not have been obvious or challenged in periods of stability but in the current fast-changing and culturally diverse societies, power relations are changing, boundaries between social practices are blurring and overlapping (Kress 2003; Cope and Kalantzis 2000; NLG 2000). In this changing landscape, it is insufficient to privilege one type of text over another. Accordingly, Kress suggests that "[a] new theory of text is essential to meet the demands of culturally plural societies in a globalising world" (2003: 120). This new theory should be "an encompassing theory of text" where genres from aesthetically valued to culturally salient and

even the banal should be included in a curriculum (Kress 2003: 120). Since literacy education aims to equip students with adequate knowledge and the right 'tools' to survive in the working world, an encompassing theory of text in literacy education is practical and in accordance with the needs of current social practices.

The second reason for choosing manga as the text of analysis is because conventions used in manga are comparable to the conventions of other visual narrative genres. Analogies between comics and films have long been noted. Research which examines the similarities between film and comic narratives has existed side by side with the structuralists' approach to comics in the 1960s (Groensteen 2000). However, the idea of using comics to study film narratives, using one visual genre to study another, is a new approach. Manga is a suitable comic genre to with which do this because it draws on conventions of film more than other types of comics. The process of using one genre to illustrate another is even more apparent in manga and storyboards as both genres are paper-based. A storyboard is "the visual version of the script. It consists of a number of panels that show the visual action of a sequence in a logical narrative" (Tumminello 2005: 1). Storyboards have often been described as film narratives illustrated in comic style. While manga and storyboards are different genres with different social purposes, the former intended for entertainment and the latter intended for production, through shared conventions, it is possible to use the one genre to illustrate the other. The similarities between the conventions of manga and the conventions of other visual narrative genres make it viable to explore the possibilities of a metalanguage for manga in interrogating other visual narratives, and subsequently contribute towards theorising a metalanguage for visual narratives.

Finally, precisely because conventions in manga are shared with other visual narratives, this particular comic genre is effective in illustrating the idea of genre as ‘designs of meaning’ (NLG 2000). In order to survive in the workplace of the twenty-first century, it is no longer adequate to be able to just replicate the rules of genres. As Kress points out,

[i]n a world of stability, the competence of reliable reproduction was not just sufficient, but the essence – on the production line as much as at the writing desk. In a world of instability, reproduction is no longer an issue: what is required now is the ability to assess what is needed in this situation now, for these conditions, these purposes, this audience – all of which will be differently configured for the next task (2003: 49).

In other words, what is needed in order to survive in a fast-changing environment is the ability to ‘design’ (Kress 2003; NLG 2000). The assumption is that semiotic resources should be seen as design resources, ‘designs of meaning’, capable of shaping and reshaping to fit the needs of the user. Because manga conventions are comparable to visual narrative genres such as film and storyboards, the text is effective in presenting the idea of conventions as design resources. This means that the conventions of one visual narrative genre can be used to reflect on the conventions of another, despite the fact that conventions are employed differently from genre to genre. The notion of genre as design resources exemplifies the weakness of genre boundaries and the social constructedness of genres.

In sum, there is a need to understand the nature of sequential visual narratives. The study contributes to this area of research by positing a metalanguage for manga. Manga is chosen as the text of analysis because of the advantages afforded through its status as a popular cultural text. In addition, the conventions of this comic genre are similar to those of other visual narrative genres. Consequently, this renders manga an ideal text in illustrating the notion of genre as ‘designs of meaning’.

By proposing a metalanguage for manga and using this metalanguage to analyse a particular manga text, the study aims to challenge the logocentric bias in narrative tradition. It challenges the belief that only the linguistic mode can adequately support the telling of a narrative by demonstrating that the choice of semiotic resources employed in manga enhances the senses and the reader's experience of the narrative rather than restricts it.

1.4 Overview of the thesis

Chapter two discusses the theoretical framework which underlines the study. It begins with an overview of our current communication landscape, describing it as 'multimodal'. The chapter then goes on to explore multimodal social semiotic theory and the implications of drawing on this theory in the study. The notion of design is proposed as a key concept in accounting for the overlapping nature of genres. The chapter argues that a metalanguage which is capable of describing different forms of meanings will be able to foreground the differences and similarities between genres. Other concepts explored are genre, medium and narrative. From a social semiotic perspective, this chapter argues that both comics and narratives are genres and that there are advantages in taking a social approach to genre theory.

Chapter three presents the methodological framework of the study. This includes a framework of the proposed metalanguage. The chapter begins with an overview of the research method and proceeds to discuss the data used in the study. Labov's (1972) narrative structure is presented as the framework outlining the data analysis. The chapter concludes with a detailed discussion of the proposed metalanguage.

Chapter four presents a detailed analysis of the data proposed for the study. The chapter is divided in five main sections. The sections correspond with Labov's (1972) proposed five events of a narrative: abstract, orientation, complicating action, evaluation and resolution. In analysing the data, this chapter addresses the first research question: "how are semiotic resources used to narrative a story in manga?"

Chapter five draws out the implications of the study. It begins with a discussion of the semiotic resources employed in the manga narrative analysed then proceeds to a discussion about modes and logics. The chapter also draws attention to the social and cultural practices in Japan and the influences these have on manga conventions. The second half of the chapter examines the possible implications of using a metalanguage of manga in interrogating other visual narratives.

Chapter Two: Theoretical Framework

2.1 Overview of chapter

This chapter presents the theoretical framework which underlines the study. It begins with an overview of multimodal social semiotics. The study argues that multimodal social semiotics accounts for the multiple modes of meaning-making in multimodal narratives and it validates all texts since all representational resources are regarded as meaningful resources. This makes it viable to bring manga into an academic context. The chapter also outlines the characteristics of an accessible metalanguage. The study argues that a metalanguage is a valuable resource as it can help identify semiotic resources and how they function to make meaning. It doing so, a metalanguage can be used to interrogate genres. Another key theory which underlines the study is a social theory of genre. The study reasons that both comics and narratives are genres and discusses the implications of adopting this view.

2.2 Multimodality and the communication landscape

In the West, the semiotic landscape of the print era is often characterised as ‘monomodal’ as it was believed that “articulate representation was by means of language” (Kress et al 2001: 2). As a result, the semiotic landscape was built around one mode – the written mode. ‘Semiotic landscape’ refers to the modes, genres, media offered by a society for communication (Kress and van Leeuwen 1996; Archer 2008). Digital technology, however, has brought about a change in perspective. Owing to the multiplicity of modes enabled by digital technology, our current semiotic landscape is now described as ‘multimodal’. There is

growing research in multimodality because the shift from the monomodal semiotic landscape of the print era to the multimodal semiotic landscape of the digital era is noted to have made drastic changes to our existence as social human beings (Cope and Kalantzis 2006; Kress 2003; NLG 2000). This shift in semiotic landscapes has radical implications for us socially, culturally, economically and intellectually because modes of human communication are understood to play an important role in shaping our thinking (episteme) and being-in-the-world (Cope and Kalantzis 2006). In other words, our epistemological and ontological assumptions of the world are shaped predominately by our communication landscape. The role of technology in shaping this new communication landscape is widely acknowledged.

Technological innovation is a prevailing force driving the revolution in the communication landscape. However, Kress (1998) argues that it would be flawed to attribute the revolution entirely to technology. According to Kress “it is both a common and a serious error to treat technology as a causal phenomenon in human, social and cultural affairs” because “[t]echnologies flourish only in part because something has become known and possible” (1998: 53). Technology is the vehicle which allows changes to take place but how the vehicle is used is subject to the inventor. In Kress’s words, “[t]echnology is socially applied knowledge, and it is social conditions which make the crucial difference in how it is applied” (1998: 52-53). Nevertheless, there is no doubt that technology played a major role in shaping our current communication landscape. In a paper titled *From Literacy to ‘Multiliteracies’: learning to mean in the new communications environment*, Cope and Kalantzis (2006) explain how technology has revolutionised modes of human communication and accordingly transformed our thinking and being-in-the-world.

According to Cope and Kalantzis, technology of the print era restricted communication largely to the written mode. This meant that the written mode became better understood as a semiotic system. It became the mode associated with knowledge and meaning. With the use of a single mode for meaning-making, it was easy to standardise and homogenise forms of meanings. In due course, meanings and ways of expression became 'fixed'. Fixing forms of meaning restricted the capacity for the negotiation of meanings, making it difficult "to recognise the role of agency to human meaning and action" (Cope and Kalantzis 2006: 29). This model of society resisted change and endorsed social hierarchies of power and order. In contrast, Cope and Kalantzis describe technology of the digital area as returning us to a world similar to that of pre-civilisation, a world of divergence, diversity and experimentation. The array of semiotic modes afforded by digital technology allow for greater negotiations and experimentations with meaning-making. Evidence of this is in the diverse texts, genres and media presently in existence. This semiotic shift is not only bringing about more means of communication but it is also changing our expectations of writing. The written mode no longer only serves the purpose of relaying a particular message but has been transformed into a "visual meaning-making resource" (Jewitt 2004: 185).

According to Jewitt, the move from page to screen is changing writing to "a visual element, a block of 'space,' which makes textual meaning beyond its written content" (2004: 185). This is certainly the case with writing in comics. Although comics typically make use of written texts in the form of speech or thought, reading in comics is ultimately a visual experience because the writing is visually oriented. As Eisner writes, "the visual treatment of words as graphic art forms is part of the vocabulary" of comics (1985: 10). In comics, writing can function as "an extension of the imagery", evoking mood through the typeface and texture employed (Eisner 1985: 10). Writing can also function onomatopoeically, evoking the sense

of sound. This means that writing is being treated more like images and like images it is subject to compositional relations. That is, writing is subject to the foreground-background continuum that is characteristic of visual design. While linguistic theories may have been sustainable in making sense of monomodal texts, such theories are no longer sufficient in explaining today's overtly multimodal texts. A new theory of communication is called for and for this reason, multimodality emerges as the new theory of communication to account for the changes in the communication landscape and the effects these changes have on modes of human existence.

2.3 Multimodal social semiotic theory

Multimodality draws on social semiotics in theorising modes of communication. In semiotics, the sign is the basic unit of meaning (Kress et al 2001). While semiotics see the sign as "an isolate, as a thing in itself, which exists first of all in and of itself before it comes to be related to other signs" (Halliday and Hasan 1985: 4), social semiotics, on the other hand, see the sign as socially oriented. In Halliday's words, social semiotics is "a social system, or a culture, as a system of meanings" (Halliday and Hasan 1985: 4). To explain this move from semiotics as 'the science of signs' to semiotics as 'a social system', it is necessary to return to the basic component of the sign.

A fundamental difference between the structuralist's approach to semiotics and social semiotics lies in their different view of the sign (Jewitt and Oyama 2001). In both semiotics and social semiotics, the sign serves as a basic unit of meaning by combining a form with a meaning (Kress et al 2001: 4). That is, the sign derives its meaning by fusing a signifier with a signified. Saussure's model of the sign is based on the premise that there is no intrinsic

relation between a signifier and a signified. The two units are arbitrarily linked. Since there is no natural link connecting a signifier with a signified, a signifier can be potentially linked with any signified. The idea of convention is therefore proposed to connect and maintain the relation. According to Saussure, “every means of expression used in society is based in principle, on collective behaviour or – what amounts to the same thing – on convention” (1966: 68). Convention binds a certain signifier to a certain signified and fixes meanings by functioning as

codes, sets of rules for connecting signs and meanings. Once two or more people have mastered the same code, it was thought, they would be able to connect the same meanings to the same sounds or graphic patterns and hence be able to understand each other (Jewitt and Oyama 2001: 134).

This means that codes have to be agreed by members of a discourse community and individuals have to learn the codes in order to use them. Therefore, from a semiotic perspective, individuals can be seen as passive users of a rigid system. In Kress’s words, “individuals are seen as users, more or less competently, of an existing, stable, static system of elements and rules” (2000: 154). Learning requires passively regurgitating conventions and success in the learner is measured by how well the learner can remember sets of conventions. Furthermore, a semiotic view of the sign suggests that change is undesirable since this would entail a relearning of the ‘codes’ again. In fact, there should be no reason for change. Since the relation between a signifier and a signified is arbitrary, it should not matter which signifier is matched with which signified. However, as Kress points out

all and any of the examples of everyday communication speak of change: changes in forms of text; in uses of language; in the communicational and representational potentials of all elements of ‘literacies’. Indeed change is one of the unchanging aspects of systems of communication (2000: 154).

The semiotic model of the sign is problematic as it assumes that the individual is passive; it does not account for relations of power that take place in social relations nor does it account

for changes that occur over time. As Jewitt and Oyama point out, “[h]ow these codes came about, who made the rules and how and why they might be changed was not considered” (2001: 135). In the print era, a semiotic approach to meaning-making was acceptable as the societies were structured on hierarchies of power. Change and questions of power were therefore undesirable. However, faced with current social instability and shifting power relations, it has become evident that a theory of social semiotics is needed to account for social factors.

Social semiotics is built on the premise that the sign is a social system of meaning (Halliday and Hasan 1985; Hodge and Kress 1988). This means that contrary to semiotics which view the sign as arbitrary, social semiotics sees the sign as always motivated. As Kress et al write, “the relation between form and meaning, signifier and signified, is never arbitrary but ...always motivated by the interests of the maker of the sign to find the best possible, the most plausible form for the expression of the meaning that (s)he wishes to express” (2001: 5). This means that a signifier is chosen for representation because of its aptness in expressing that which the individual wishes to mean rather than for arbitrary reasons (Kress 2000, 2003).

According to Kress, the use of particular signs to express a particular meaning is “an effect both of the demands of particular occasions of interaction and of the social and cultural characteristics of the individual maker of signs” (2000: 156). This suggests that the sign is motivated by two factors. Firstly, the interest of the sign-maker in representing a phenomenon in a particular context and secondly, the socio-cultural trends associated with using particular signs (Kress 2000). From this perspective, signs are not a system of ‘codes’ but rather a system of ‘resources’ which an individual draws on for expression. The idea of signs as ‘resources’ rather than ‘codes’ is a vital distinction between social semiotics and

semiotics (Jewitt and Oyama 2001). A system of signs built on conventions that are arbitrarily constructed between members of a discourse community suggests that there is no purpose in studying signifying practices since conventions are meaningless codes. Furthermore, the view of conventions as codes means that change is not encouraged. According to Kress, “the common understanding is that convention impedes change, that convention is a force for the maintenance of stability” (2000: 154). However, a system of signs built on the notion of *social convention*, social in the sense that they are motivated by social factors such as the interests of the discourse community that built them and the socio-cultural trends of the particular time of construction, means that sign practices will change with social changes and studying signs will provide an understanding of human signifying practices. As Kress points out, a social semiotics approach to signs

means that signs are always meaningful conjunctions of signifiers and signifieds; it means that we can look at the signifiers and make hypotheses about what they might be signifying in any one instance, because we know that the form chosen was the most apt expression of that which was to be signified...It entails that all aspects of form are meaningful, and that all aspects of form must be read with equal care; nothing can be disregarded (2003: 44).

The implications of a multimodal social semiotic theory are important for this study. For one, a social semiotic approach to multimodality means that all semiotic resources are meaningful. Since the study aims to examine the affordances of various semiotic resources in storytelling, this premise is central to the study. The term ‘affordance’ refers to “the potential uses of a given object” (van Leeuwen 2005: 4). The idea of signs as resources suggests that meanings can be negotiated – they can be designed and manipulated to suit the interests of the user. This suggests that genre conventions can be renegotiated, redesigned, essentially reshaped. The variability of genre conventions speaks of fluid genre boundaries. What this means is that it becomes viable to see “how the [genres] you already have relate to those you are attempting to acquire, and how the ones you are trying to acquire relate to self and society”

(Gee in Thesen 2001: 143). A metalanguage of one genre can then serve as “an index of discourse” (Thesen 2001: 143) for a number of genres. The implication is that one could use a metalanguage from one genre to interrogate another. Multimodal social semiotics provide an explanation for the changes that happen in our social environment and it affords individuals far greater agency than past theories of semiosis. This opens doors for different ways of learning and validates more genres of texts for use in the classroom.

2.4 A multimodal social semiotic approach to genre

The New London Group (NLG) argues that the primary purpose of education is “to ensure that all students benefit from learning in ways that allow them to participate fully in public, community and economic life” (2000: 9). With the aid of new media technologies, these spheres of our social lives are characterised by multiplicity, diversity and change. Our workplace, for instance, is characterised by a post-Fordist, fast capitalist work culture (NLG 2000). What this means is that “old hierarchical command structures” are being replaced by a “flattened hierarchy” and “mindless, repetitive unskilled work” is being replaced by work which requires “multiskill[s]” (NLG 2000: 9). Innovation and creativity are therefore key skills to surviving in such a work environment. Our social lives are characterised by diversity as globalisation brings together different cultures and societies. Cultural and linguistic diversities not only affect our social lives but also our working lives. It is thus important that education provide students the skills to negotiate these differences. The advent of new media technologies also means that a plurality of texts, genres and modes are on the increase. This means that students need a theory of literacy which can account for the plurality of meanings in these texts.

A genre approach to literacy involves foregrounding genre conventions by explicitly teaching the grammars of a language (Cope and Kalantzis 1993; Kress 1993). The purpose for highlighting genre conventions is to show the social constructs behind the construction of a genre – that is “show what kinds of social situations produce them, and what the meanings of those social situations are” (Kress 1993: 24). By drawing attention to the conventions of a genre, the genre approach hopes to establish “a sufficient understanding of grammar as a dynamic resource for making meaning” (Kress 1993: 24).

There are two underlying principles supporting a genre approach to literacy. Firstly, it seeks to “establish[_] a dialogue between the culture and the discourse of institutionalised schooling, and the cultures and discourses of students” by highlighting how genre conventions work in particular contexts to produce particular meanings (Cope and Kalantzis 1993: 17). A metalanguage to make generalisations about how language functions can enable students to see the similarities and differences between academic genres and genres-in-use. A genre approach to literacy is not just about the teaching of rules but rather about how conventions work to achieve their social goals. Secondly, genre literacy seeks “to provide historically marginalised groups equitable access to as broad a range of social options as possible” (Cope and Kalantzis 1993: 8). Genre theorists argue that by making conventions of academic genres explicit, by denaturalising genres of power, this will allow learners outside these discourses potential access.

Despite the noble intentions of genre literacy to provide a fair and viable approach to literacy, Luke warns the risk of “renaturalising” genres of power and reaffirming the status quo of the ruling class by failing “to situate, critique, interrogate, and transform these texts, their discourse and their institutional sites” (1996: 334). Luke argues that pedagogies should go

beyond the teaching of rules of a genre and focus on the “social and cultural strategies for analysing and engaging with the conversion of capital in various cultural fields” (1996: 332). Kress echoes this notion and suggests that “a newer way of thinking may be that within a general awareness of the range of genres, of their shapes and their contexts, speakers and writers newly make the generic forms out of available resources” (2003: 121). That is to say, genre conventions should be seen as design resources that can be reshaped to an individual’s needs. For this reason, Kress (2003) proposes a multimodal view of genre and an approach to literacy pedagogy where texts from all realms of social practices, from the aesthetically valued to the culturally salient to the banal texts of everyday, are encompassed in a curriculum. By drawing on genres from various social realms for analysis, this will create an open dialogue between academic genres and genres-in-use. It will visibly illustrate the idea of genre as social strategies for achieving particular social purposes and accordingly the idea of convention as design resources for making meanings will become apparent.

The notion of viewing conventions as design resources is particularly crucial in our current social environment. Boundaries that separated social practices are no longer as clear as they were in the print era. Social practices are overlapping with one another and with this, the mixing of genres is becoming more and more evident. In this environment, teaching rules and restricting the school curriculum to particular genres is impractical because this does not reflect the type of skills needed to succeed. Bourdieu’s concept of ‘symbolic capital’ points to the fact that ultimately, capital and power are only valid if they are valued by society as such. As Luke writes, “ultimately, capital is only capital if it recognized as such; that is, if it is granted legitimacy, symbolic capital, within a larger social and cultural field” (1996: 329). Bourdieu’s theory of ‘symbolic capital’ suggests that “it is misleading to assume that any genre, skill, text has generalisable power, tied up with a singular kind of capital in social

structures” (Luke 1996: 330). This is particularly evident in our current environment where new genres and texts are constantly emerging. Thus to keep literacy education valid, it is necessary to have a theory of genre which recognises the shifting power relations behind social practices, which recognises the validity of all genres and which allows for the mixing of genres.

A multimodal approach is non-mode and non-genre specific; modes and genres are seen as motivated design resources. All texts are valuable as different texts serve different functions. A multimodal view of genre is one that is interested in understanding “what is it that we want to mean, and what modes and genres are best for realising that meaning” (Kress 2003: 107). As an approach to pedagogy, multimodal pedagogies foreground the affordances of various modes and seek to understand how different modes function to produce different forms of meanings in particular contexts (Kress et al 2001; Archer 2006b). Metalanguages assist multimodal pedagogies by functioning as tools of analysis. They act as “languages of reflective generalization that describe the form, content and function of the discourses of practice” (NLG 2000: 34). A multimodal approach to genre has the potential to provide students with strategies to engage in conversions of capital and access to new realms of social practices. It is, as Kress writes, “a much more ‘generative’ notion of genre: not one where you learn the shapes of existing kinds of text alone, in order to replicate them, but where you learn the generative rules of the constitution of generic form within the power structures of a society” (2003: 121). This study explores a multimodal view of genre by constructing a metalanguage from a popular cultural text and examines the implications of using this metalanguage in interrogating other visual narrative genres.

2.5 Characteristics of an accessible metalanguage

The aim of this study is to develop a metalanguage of analysis for manga and to explore the possible implications of using this metalanguage in interrogating other visual narratives. A metalanguage is a set of grammars, to describe how meaning is produced in various modes. It is a language to make “generalisations” about semiotic resources (Cope and Kalantzis 1993: 8) so that it is possible to describe and understand how they function to produce meaning in particular contexts. This is in line with social semiotics where semiotic resources derive their meaning from their context of use. According to the NLG (2000), a metalanguage to be used in a classroom context should possess three particular qualities.

Firstly, a metalanguage must be adequately developed so that it allows for critical analysis but “at the same time not make unrealistic demands on teacher and learner knowledge” (NLG 2000: 24). This suggests that while a well-developed metalanguage is called for, it must also be a workable framework, practical for teachers and learners alike. The metalanguage proposed here is based on Kress and van Leeuwen’s (1996) extended framework of Halliday’s systemic functional linguistics and social semiotics. Kress and van Leeuwen’s metalanguage has proved to be very successful in the analysis of still images. To account for the sequential nature of images in comics, this study draws on Matthiessen’s (2007) work on image-image relations. The two frameworks are described in greater detail in chapter three. Although both frameworks are extremely useful, they can be somewhat complex for learners who have not previously encountered a grammar for images. Particularly in South Africa, where English is not the first language of many students, such a complex framework can be daunting and may hinder rather than enhance the learning experience. Thesen suggests that rather than propose an entirely new set of vocabularies and new ways of understanding the

world, a practical metalanguage should serve as “an index of discourse – ways of verbalising what you know in relation to other ways of knowing” (2001: 143). Hence, to provide a functional metalanguage for students, Matthiessen’s and Kress and van Leeuwen’s frameworks have been modified so that they can be accessible to students.

Secondly, “a metalanguage also needs to be quite flexible and open ended” (NLG 2000: 24). A metalanguage is not meant to be viewed as a set of rigid rules which should be applied. Rather, the NLG suggests that it should be viewed as a ‘toolkit’ – “[t]eachers and learners should be able to pick and choose from the tools offered. They should also feel free to fashion their own tools” (2000: 24). Texts change with social changes. Therefore, a flexible metalanguage is required to cope with changes. The idea of a flexible metalanguage and a metalanguage as a ‘toolkit’ is particularly important for this study. The metalanguage proposed here is constructed around the analysis of comics which means that when interrogating other visual narrative genres, there will be aspects of the metalanguage that will not apply. The idea of metalanguage as a ‘toolkit’ rather than a set of static rules foregrounds the idea of a flexible metalanguage that can be transformed to suit the situation of analysis.

Finally, a metalanguage should function to “identify and explain difference between texts, and relate these to the contexts of culture and situation in which they seem to work” (NLG 2000: 24). In other words, it is necessary for a metalanguage to provide strategies in critically engaging with texts, in identifying the social purposes, the genre of a text, and how they function to attain that goal.

A metalanguage is important because it allows for critical analysis of modes and creates a sufficient understanding of the affordances of various modes. A metalanguage for describing

modes of communication other than language has been non-existent in the past and consequently, this has hindered our understanding of their potentials as resources for making meaning. The Sapir-Whorf theory proposes that people's worldviews are shaped by the grammatical structures of the language they use (Chandler 2007). While most contemporary critics agree this thesis is somewhat extreme, it nonetheless speaks of the critical role language plays in shaping people's worldviews. Barthes, for example, favoured a logocentric view of the world. According to Barthes,

It is true that objects, images and patterns of behaviour can signify, and do so on a large scale, but never autonomously; every semiological system has its linguistic admixture. Where there is a visual substance, for example, the meaning is confirmed by being duplicated in a linguistic message...so that at least a part of the iconic message is, in terms of structural relationship, either redundant or taken up by the linguistic system. As for collections of objects (clothes, food), they enjoy the status of systems only in so far as they pass through the relay of language...there is no meaning which is not designated, and the world of signified is none other than that of language (1967: 10).

Barthes' logocentric bias could be partly attributed to the fact that at the time of his writing, despite the wide spread use of visuals, there were few metalanguages available to discuss the affordances of other semiotic modes. After all, it is difficult to understand, much less, talk about the meaning-making potentials of a semiotic system without a metalanguage to describe it. For this reason, over the last few years educators have been working at developing and providing metalanguages for modes in various realms. A metalanguage for visual narrative genres such as comics and picture books has yet to be developed. With the growing popularity of visual narratives, it is evident that such a metalanguage is necessary. The purpose of this study is to contribute towards theorising a multimodal pedagogy by developing a metalanguage for manga and looking at the possibilities of using this metalanguage in interrogating other visual narratives.

2.6 Genre as ‘designs’

The concept of design is important when looking at the possible implications of using the proposed metalanguage in interrogating other visual narratives. According to the NLG (2000), the notion of design is based on a theory of discourse which recognises knowledge and meaning as socially, culturally and historically constructed truths; they are ‘design artefacts’. Semiotic resources are thus seen as design resources and meaning-making as an active and subjective process. This concept aligns with a theory of social semiotics which posits the sign as motivated by the interests of the sign-maker and by the socio-cultural context in which the individual is entrenched (Kress 2000; Kress et al 2001; Kress 2003). A pedagogy based on design accounts for change as semiotic systems are viewed as “dynamic, constantly remade and reorganised set of semiotic resources” (Kress 2000: 157). This is because the concept of design sees any semiotic activity as involving three design elements: Available Designs, Designing and The Redesigned (NLG 2000). Available Designs refer to resources that are available for meaning-making. Designing is the process where Available Designs are reworked to produced new resource for meaning-making. This is the transformation phase as resources are changed into something new. The Redesigned is the product of the Designing process. The NLG posits that to fully support a pedagogy of design, a metalanguage is necessary to described the forms of meaning represented in the Available Design and The Designed. This means that by highlighting the similarities and differences between the design resources, it brings out the possibility of using one genre to interrogate another.

According to the NLG, “[t]he notion of Design recognizes the iterative nature of meaning-making, drawing on Available Designs to create patterns of meaning that are more or less

predictable in their contexts. This is why *The Redesign* has a ring of familiarity to it” (2000: 22). This suggests that when one genre of text is transformed into another there is always a trace of the old in the new. The NLG adds that “*The Redesign* is founded on historically and culturally received patterns of meaning” (2000: 22). This points to the fact that there will always be hybridity and intertextuality in genres. Certainly, the group describe genre as “an intertextual aspect of a text”. According to the NLG, genre “shows how the text links to other texts in the intertextual context, and how it might be similar to other texts used in comparable social contexts, and its connections with text types in the order(s) of discourse” (2000: 25). A metalanguage which can describe the different forms of meaning in genres is therefore a valuable resource in enabling one to see the inter-relationship between genres.

2.7 Comics from a social semiotic perspective

Meaning-making in comics depends on the integration of a number of semiotic resources. In comic studies, there have been continuous debates over what texts can be considered as comics and whether comics are a medium or a genre. The two debates are in fact inter-related and they both stem from a lack of a unified definition of comics. From a social semiotic perspective, ‘medium’ and ‘genre’ have different properties and different affordances. For this reason, it is important to outline the properties of comics and establish a definition for this study.

Will Eisner’s (1985) definition of comics as ‘sequential art’ has been by far the most influential and supported definition to date. By ‘sequential art’ Eisner means “the arrangement of pictures or images and words to narrate a story or dramatize an idea” (1985:

5). Most comic critics agree on this definition probably because ‘sequential art’ contains the bare necessities of what should be considered as comics and at the same time it is vague enough for open interpretation. For example, the concept of sequential art can be narrowed down to specific types of texts such as in the case with Sabin’s definition. According to Sabin,

[t]he fundamental ingredient of a comic is the ‘comic strip’. This is a narrative in the form of a sequence of pictures – usually, but not always, with text. In length it can be anything from a single image upwards, with some strips containing images in the thousands. A ‘comic’ per se is a publication in booklet, tabloid, magazine or book form that includes as a major feature the presence of one or more strips (1993: 5).

From Sabin’s perspective, comics as sequential art can only happen in the form of books, newspapers and magazines. Contrary to Sabin, McCloud (1994) interprets Eisner’s idea of sequential art to be inclusive of any texts as long as the images in the text follow a sequence:

com-ics (kom'isk) **n.** plural in form, used with a singular verb. **1.** Juxtaposed pictorial and other images in deliberate sequence, intended to convey information and / or to produce an aesthetic response in the viewer (McCloud 1994: 9).

This definition has been most controversial largely because it is broad and allows a number of texts not conventionally considered as comics to fall into the category. In McCloud’s words, “[f]rom stained glass windows showing biblical scenes in order to Monet’s series paintings, to your car owner’s manual, comics turn up all over when sequential art is employed as a definition” (1994: 20). Among all the texts McCloud lists as comics, the Bayeux Tapestry, a two hundred and thirty foot long tapestry featuring the eleventh century Norman invasion of England, has been the most controversial. Sabin argues that by including the tapestry into the category of comics, McCloud’s definition is deliberately vague for political reasons. Sabin reasons that the type of cultural status associated with the Bayeux Tapestry contrasts greatly with the “despised art form, barred from serious critical discussion

and stereotyped as either kids' stuff or as a pastime for nerds" (2000: 48) that has traditionally been affiliated with comics, particularly those from the US and UK.

The fundamental reasoning behind McCloud's broad definition is to be able to present comics as a medium. McCloud argues that comics should be understood as "a vessel which can hold any number of ideas and images" and should not be confused with the contents presented through the medium (1994: 6). In other words, comics should be understood as a medium and not as a genre. McCloud would want to posit this view for political reasons.

Genre is often associated with repetitions, formulas, clichés, features considered as generally negative. However, this negativity was not always present. Kress and Threadgold (1988) observe that in classical periods genre was a valued term and all that was considered literature had to be generic. This concept owes itself to the association of rules with being polite and civilised. In the nineteenth century, this view altered along with social and cultural changes brought by industrialisation. Rules came to mean constraint and lack of creativity. So, by the twentieth century when genre spread to the classification of popular cultural texts, it became more of a negative concept associated with being stereotypical (Kress 2003).

Modern comics emerged in this period of industrialisation. The advent of the printing press made it easy to mass produce comics for a mass audience (Sabin 1996). Subsequently, comics have been labelled as a genre for the masses and any genre associated with the masses is generally perceived as 'low art'. Groensteen points out that "for the educators of the first half of the twentieth century, that which is popular is necessarily vulgar" (2000: 32). Beyond questions of culture and aesthetics, there are other reasons why comics have, in the past, received negative reviews from educators. In a chapter entitled *Why are comics still in search*

of cultural legitimisation, Groensteen notes that comics seem to be “condemned to artistic insignificance” because of a “four-fold symbolic handicap” (2000: 35):

1) It is a hybrid, the result of crossbreeding between text and image; 2) Its story-telling ambitions seem to remain on the level of a sub-literature; 3) It has connections to a common and inferior branch of visual art, that of caricature; 4) Even though they are now frequently intended for adults, comic propose nothing other than a return to childhood (Groensteen 2000: 35).

The first three of these ‘handicaps’ notably points to the unease with the hybrid nature of comics. Kress (2003) notes that in the history of the West, much emphasis has been placed on maintaining social control and establishing social stability. Hybridisation causes unease because it is about social flux. According to Groensteen,

Comics are seen as intrinsically bad and because they tend to take the place of ‘real books’, an attitude which crystallizes a double confrontation: between the written world and the world of images, on the one hand; between educational literature and pure entertainment on the other (2000: 32).

The straddling between various spheres, crossing between two modes of expression (language and images) and two genres (education and entertainment), presents comics as unstable and ambiguous. In the West, that which is ambiguous is generally treated as a threat, a taboo (Douglas 2005). According to Douglas, taboo is the “spontaneous device for protecting the distinctive categories of the universe. [It] protects the local consensus on how the world is organised” (2005: xi). Comics are often met with disapproval largely because their hybrid nature threatens social stability. Furthermore, as in any social practice, there are power relations involved.

Certain genres are seen as more valuable because they are associated with social practices which afford power. For Bourdieu, power emerges in the form of various types of ‘capital’. Capital range from political, to economic, to cultural or symbolic. According to Bourdieu

(1990), “capital, like trumps in a game of cards, are powers which define the changes of profit in a given field” (cited in Luke 1996: 326). Hence, social practices are never ‘neutral’ but they “take place in fields of power” (Luke 1996: 326). It follows that “genres which are characteristic of a social group are not just expressions of such power, they are also arranged in hierarchies of power” (Kress 2003: 85). Groensteen argues that one of the ‘handicaps’ of comics is their association with children’s reading practices. This is not a practice that is seen to afford power.

In sum, comics suffer from several ‘handicaps’ in the eye of educators because this genre neither abides to the laws of ‘purity’ that govern Western literary discourse nor are they a genre associated with powerful practices in society. With all the negative connotations discussed above, by positing comics as a medium, McCloud avoids having to justify comics as a genre and focuses purely on the representational potentials of comics. Moreover, the term ‘medium’ can lend the notion of stability that is lacking in discussions surrounding comics as a genre.

According to Kress and van Leeuwen, a medium is the “material resources used in the production of semiotic products and events” and it serves the purpose of recording and distributing the semiotic products and events produced (2001: 22). As such, a medium acts like a structure, a frame which holds genres of various kinds. For example, a book is a medium which can accommodate a range of genres such as poetry, drama and romance. Likewise, television is a medium which can broadcast a number of genres ranging from documentaries to soap operas. In this sense, the concept of medium is ‘larger’ than genre. Rommens reflects a similar thought when he writes “Western discourse often ‘annexes’ manga in the overall European/American comics production by representing it as a mere

genre within comics' constellation, thereby denying the fact that manga is a medium in its own right" (2000). In using the word "mere" to describe genre, we are presented with the idea that genre is something inferior. In contrast, the words "in its own right" hint at dignity in the term 'medium'. By equating comics to a medium, McCloud attempts to elevate the position of comics and portray them as a stable entity which demands respect. This notion is reinforced by presenting his definition in a convention commonly associated with dictionaries.

Yet, despite McCloud's efforts in positing comics as a medium, from a social semiotic perspective, the term 'medium' does not adequately describe comics. In Kress and van Leeuwen's (2001) definition of a medium, they emphasise two fundamental qualities. The first is materiality – a physical quality. As mentioned earlier, a medium comprises "the material resources used in the production of semiotic products and events, including both the tools and the materials used (e.g. the musical instrument and air; the chisel and the block of wood)" (Kress and van Leeuwen 2001: 22). A book, for example, is a medium. When it functions to communicate a message, it does so mainly through paper and ink. The second quality is related to the first. As a physical, material entity, a medium should have the ability to record and to distribute. A medium is "developed specifically for the recording and/or distribution of semiotic products and events" (Kress and van Leeuwen 2001: 22). In this sense, no matter what the content is or what modes are used in creating the text, the properties of the medium will not change. Whether we read Shakespeare's *Hamlet* with language as the mode of communication or Frank Miller's *Batman* with images as the primary mode of communication, the two very different texts still function through the same medium. The materiality of the resources used in both texts is still associated with books. Likewise, in television, no matter what programme is broadcast, whether it be a documentary

program or a soap opera, the audience still receives the text through the medium of a television.

If we accept the definition of medium described by Kress and van Leeuwen, then the concept of medium is less appropriate to describe comics. As McCloud strongly argues, comics can grace a number of mediums ranging from tapestry to the Internet. This means that comics are not bound to the material resources of their production. While the sequential nature of comics allows them to communicate in a unique manner, without the material resources of the medium they appear in, they cannot function to record or to distribute. Therefore, this study contends that from a social semiotic perspective, the term ‘medium’ does not adequately describe comics. When McCloud argues that comics are a medium because like other media, they are “a vessel which can hold any number of ideas and images” (1994: 6), it is possible that he may be mistaking a genre convention for a medium.

Genre is a principle of textual organisation which functions “to make *form* (the conventions of the genre) more transparent to those familiar with the genre, foregrounding the distinctive *content* of individual texts” (Chandler 2007: 189). Consequently, there is a tendency to identify a genre either through its formal properties or content. Formal properties can be understood as the particular structures, grammar and semiotic resources which allow the genre to communicate its message in a particular way. McCloud claims that comics are a medium because various contents can appear in the art form. However, formal properties of a genre are also capable of fulfilling this function. Emails, for example, are a genre of letter writing with distinct formal properties clearly identifiable as email conventions. These properties include the address line, the forward line, the subject line and the writing box. Like comics, emails can ‘hold’ a variety of content through their formal properties. Panels, speech

and thought bubbles are formal properties of comics and while various contents can surface through these properties, this does not qualify comics as a medium for the reasons mentioned above.

As seen, despite varying notions of what comics are, all comic critics agree on the sequential nature of comics. A reason why critics have trouble defining comics can be a result of their social nature as a genre. This study argues that a social theory of genre can adequately account for the varying comic conventions found across societies and historical periods and therefore posits the view of comics as a genre. Perhaps provoked by McCloud's take on his notion of 'sequential art', Eisner later revises his definition and concludes that comics are "the printed arrangement of art and balloons in sequence particularly as in comic books" (1996: 5). By drawing on conventional features of modern day comics, this definition is useful for understanding what comics are today. In short, this study sees comics as a visual narrative genre. That is, comics are a visual genre with images employed in sequence for narrative purposes. Modern conventional features of this genre include the use of frames, speech and thought bubbles. Comics generally appear in the medium of print. However, as a genre, they are medium independent. The next section will discuss the benefits of taking a genre approach to comics.

2.8 A social theory of genre

The term 'genre' comes from French meaning 'type' or 'kind' (Neale 2000). It was initially used to describe distinct literary forms in literature but it has since become a broad concept, understood as a principle of categorisation, as a basis for text distinctions. As mentioned above, genre as a principle of categorisation has tended to classify texts either by their formal

properties or content. However, Chandler points out that this view is “deeply problematic” as it “ignores the way in which genres are involved in a constant process of change” and it does not take into the account that “genres overlap and texts often exhibit the conventions of more than one genre” (2007: 158). For this reason, more recent theories of genre have tended to adopt a social approach (Cope and Kalantzis 1993; Kress 1993; Kress 2003; Luke 1996). A social approach to genre is effective in accounting for the changing and hybrid nature of genres.

Social semiotic theory regards genre conventions as socially constituted. Conventions which make up a genre are not a given or decided by an individual alone. In order to distinguish one genre from another, there has to be mutual acceptance and understanding among members of a specific group that certain features (through repetitive use) constitute a convention for a specific genre. For this reason, Hodge and Kress note that, “genres only exist in so far as a social group declares and enforces the rules that constitute them” (1988: 7). If genres are socially constructed then like all social practices, genres are subject to social, cultural and historical factors. This is a crucial point. Social, cultural and historical factors change over time. If genres are subject to these factors then genres too must be constantly changing. This point is important as it explains why there are varying conventions in comic genres. As a social practice, comic conventions evolve over time. For example, speech and thought bubbles which are characteristic of modern comics were not conventions of earlier forms of comics (Sabin 1996). In addition, because genres are socially constructed, this suggests that they are immersed in their social contexts of use. This point explains the reason why manga displays particular conventions which make them distinct from Western forms of comics, despite being a comic genre. Such conventions include reading from right to left, exaggerated facial expressions and body gestures of characters and the use of emoticons (Rubinstein-

Ávila and Schwartz 2006; Kinsella 2000). These conventions arise from their context of situation. Therefore, although genres are recognisable through certain settled conventions, as social practices the conventions are always involved in a constant process of change, subject to social, cultural and historical factors. This leads to another important feature of genre. Since genres are constantly changing and evolving with social changes, this means that they are hybrid in nature.

According to Kress, “[t]he mixing of genre has to be a reality, simply as an effect of our ordinary normal social lives and our ordinary normal use of language; constant change has to be seen as entirely normal as an effect of a social theory” (2003: 87). Kress argues that the mixing of genres is unavoidable because as in any social process, there will always be new situations which requires new genres. New genres, however, are never entirely new. They are always built on the genres before but changed in particular ways to suit the context of use. While the hybrid nature of genres may not have been so apparent in the past, today hybrid texts are particularly evident as societies negotiate and experiment with various ways of making meanings. The mixing of genres is also incited by boundaries between communities of practice collapsing due to “social pressures and to changes in their own institutional, professional and organisational structures, or simply because of the sheer accretion of knowledge” (Candlin and Sarangi in Bhatia 2004: x). This can be seen in comic conventions. Images in manga are framed and drawn in a similar manner to film genres. One particular practice unique to manga is the use of ‘subjective motion’. According to McCloud (1994), this technique places the reader in the narrative by allowing the reader to see images from the character’s perspective. This is a technique often employed in films and “operates on the assumption that if observing a moving object can be involving, being that object should be more so” (McCloud 1994: 114). With filmic images becoming the norm as a result of

television, film and photographs, it comes as no surprise that manga would draw on conventions of film to match the taste of modern audiences.

The hybrid nature of genres can also be attributed to the intertextual nature of all texts. 'Intertextuality' is a term coined by Julia Kristeva and it refers to the relationship a given text may have in relation to others (Chandler 2007). Fairclough (1992) distinguishes between two types of intertextuality: manifest and constitutive. Manifest intertextuality refers to the inter-relationships between the content or context of texts. It is the "heterogenous constitution of texts out of specific other texts" (Fairclough 1992: 85). This means that the contents or contexts being drawn on are explicitly present. Parody and satire are forms of manifest intertextuality. They operate on the basis that the reader is familiar with the content or context of the situation. Constitutive intertextuality or interdiscursivity refers to the inter-relationship of types of discourse such as genres. A genre can draw on discursive features of a number of genres. Manga, for example, draws on conventions of both film and comics. According to Fairclough, "[t]he concept of intertextuality points to the productivity of texts, to how texts can transform prior texts and restructure existing conventions (genres, discourses) to generate new ones" (1992: 102). It can thus be said that the intertextual nature of texts result in hybridisation of genres.

A social theory of genre is a powerful way of connecting texts and social practices. As Kress writes, genre is "one aspect of textual organisation, namely that which realises and allows us to understand the social relations of the participants in the making, the reception and the reading/interpretation of the text" (2003: 94). Knowing who the participants in a genre are and the nature of their social relations with one another will lead to an understanding of the kinds of situation that produce the genre and the meanings of those social situations. This

means that by taking a genre approach to analysing a text, we will not only know what the purpose of a genre is but also how conventions are staged to achieve that purpose. As Martin and Rose note, “genre is a staged, goal-oriented social process” (2003: 7). This suggests that genre conventions are less like static rules and more like flexible designs. Many genres draw on the same conventions but how the conventions are employed or designed differ from genre to genre because each has its own social purpose. For example, manga and storyboards are two separate genres but they share conventions. Both of these genres make use of framed images in sequence and both of them draw on conventions of film discourse. However, the two genres have different social purposes and therefore employ their conventions differently. Comics aim to entertain thus conventions are designed to provide a smooth reading experience. To encapsulate readers in the narrative, film conventions are implicit in the images and speech and thought bubbles are employed to keep dialogues and thoughts within the narrative world. In contrast, storyboards aim to instruct. A storyboard functions to provide as much information as possible for directors or cinematographers to frame and capture the desired image. This means that film conventions are explicitly spelled out in writing and the use of diagrammatic arrows indicating action are common. Dialogues are written outside the framed images and they exist more for reference than for narrative purposes. Therefore, storyboard conventions are designed to give appropriate instructions while comic conventions are designed to entertain. Although, there are similarities between comic and storyboard conventions, they are employed differently because the two genres have different social purposes.

This example illustrates that by studying a text from a genre approach, it can provide insight into how conventions are employed to achieve the goals of a text. It also illustrates that genre conventions are like design resources that can be shaped accordingly. From a pedagogical

perspective, this view of genre suggests that it is viable to use a metalanguage from one genre to interrogate another.

In summary, a social theory of genre accounts for the variations of conventions across genres as well as the overlapping nature of genres. It also draws attention to the social constructedness of genres. In doing so, genre is presented as a “*social strategy* historically located in a network of power relations in particular *institutional sites* and *cultural fields*” (Luke 1996: 333). By drawing attention to the social constructedness of genres, it demystifies the authority associated with particular genres. By presenting genre as a social strategy and by drawing attention to the similarities and differences of one genre with another, it highlights the notion of genre as designs. This proposes the possibility of learning new genres in relation to what is already known. In other words, it points to the possibility of converting capital. This study aims to demonstrate the validity of this premise by constructing a metalanguage for manga and applying this metalanguage to analysing other visual narrative genres.

2.9 Narrative as a genre

Narrative is a popular topic of inquiry for a number of disciplinary fields. A definition of narrative remains in dispute as different theoretical approaches offer different definitions. Despite this, most theorists agree that a narrative is composed of two structures: a story and a plot. ‘Story’ refers to the sequence of events that are represented. It is concerned with the objects and actions that make up a story. ‘Plot’, on the other hand, refers to the presentation of the story. It is the mediation of the story through a storyteller. Various terms have been used to describe these two structures. Russian formalists such as Propp referred to them as

fabula and *sjuzhet*, while French structuralist, Barthes, used the terms *histoire* and *discourse* (Toolan 1988). Chatman (1978) made the distinction between story and discourse. This study uses the terms story and plot.

The story component is concerned with the happenings of a sequence of events. An event is usually perceived as a social activity. Social activities always involve actors or characters of some kind. They also always occur in particular contexts or settings. An event therefore must comprise characters and settings. Chatman (1978) uses the term 'existents' to describe these elements. In addition, actions involved in any event. Characters take on particular roles and as a result of their actions, something always 'happens' to the characters. This notion of actions and consequences is further pronounced when the story is viewed as a whole, not of just one event but as a sequence of events. Chatman uses the term 'events' to describe actions and happenings. According to Chatman, viewed as a whole, a story consists of "the content or chain of events (actions, happenings), plus what may be called the existents (characters, items of setting)" (1978: 19). It must be noted that there is a time dimension involved in a story. Events always happen in time. The notion of a sequence of events itself evokes the notion of progression in time. Time in the story component of a narrative always follows the natural order of events. However, this time can be manipulated in the re-presentation of the narrative. Plot is that component of a narrative structure which is involved with the manipulation of the 'natural' order of events of a story.

Plot describes the component of a narrative which is concerned with the construction of a story, the re-presentation of a story to an audience. Chatman uses the word 'discourse' and suggests that discourse is the "the expression, the means by which the content is communicated" (1978: 19). This suggests that 'discourse' is the order of the story that is

presented *plus* the manner and medium in which the narrative is presented. Every act of communication necessarily requires a means of expression and a medium of communication. This means that ‘discourse’, as Chatman uses it, is not a feature unique to narrative. For this reason, this study works with the concept of ‘plot’ and understands plot to mean the structuring of the story. The plot is that component of a narrative which is involved with the manipulation of the happenings of the story. This includes the manipulation of time (order and duration) and perspective (point of view).

The terms ‘genre’ and ‘discourse’ have both been used to describe narrative. This study sees genre as an aspect of textual organisation that is socially constructed to carry out certain communicative functions. They are characteristic of relatively stable conventions. However, being socially situated, conventions of a genre do change with changes in the social environment. On the other hand, this study understands discourse to mean “a way of signifying a particular domain of social practice from a particular perspective” (Fairclough 1995: 14). That is, discourse is the use of language to carry out particular activities and identities associated with particular social groups. Gee refers to this as “language-in-use” (1999: 7). Discourse is intrinsically linked to an individual’s social identity, ideology and beliefs. Bourdieu’s concept of ‘habitus’ suggests that different individuals will draw upon different discourses when communicating because the discourses an individual draws on are subject to an individual’s socio-cultural and historical background. In contrast, genre is a social practice that any individual from different discourse communities can draw on to enact particular communication purposes. For example, two students are asked to write an essay on whether they think abortion should be legalised or not. The one student is a female and a human rights activist. The other student is a male and a Christian. Since this is an essay, both will draw upon the essay genre to write their topic. However, the discourses they draw upon

to make their arguments will differ. The male student could possibly draw upon religious discourse to argue that abortion is wrong. The female student, on the other hand, could draw on human right discourse. The discourses they draw on are subject to their socio-cultural, historical and even their social positions in society as female and male. From this perspective, the term 'genre' better describes narrative than 'discourse'.

If we view narrative in terms of story and plot, then narrative is clearly a social practice with distinctive structures. Furthermore, according to Branigan, "[m]aking a narrative is a strategy for making our world of experiences and desires intelligible. It is a fundamental way of organising data" (1992: 1). This concept is widely accepted and it points to the fact that narrative has a functional purpose. Looking at how narratives of personal experience function within our society, sociolinguist William Labov identifies two particular functions of narrative. On the one hand, narrative serves the function of "recapitulating" experience (1972: 359). That is, it has the function of summarising experience. On the other hand, narrative serves the function of evaluating an experience. This suggests that narrative has a communicative purpose rather than an ideological purpose. Narrative structures allow one to organise aspects of a text in order to recapitulate and evaluate experience. From this perspective, narrative is better thought of as a genre than as a discourse.

Most theorists note that narrative is a universal phenomenon. Narratives are all pervasive, found in every culture and every society. Even though approaches to constructing a narrative differ from place to place and practice to practice, it is mostly recognised as an approach to organising text for particular communicative purposes. Narrative is not specific to a social group although different discourses do emerge in the telling of a narrative. This is because every act of communication is constituted through discourse. The same narrative can be told

from different perspectives through various media and yet it is still recognised as a narrative. While narratives may be recounted differently across cultures and societies, the concept of narrative is not specific to any culture or society. Discourse thus does not adequately account for the concept of narrative. However, a social theory of genre serves well in describing narrative as an aspect of textual organisation, a particular convention for communication, and also in accounting for the varying types of narratives that can be found.

2.10 Final Comments

The aim of this study is to take a tentative step in proposing a metalanguage for visual narratives by looking at a particular visual narrative text, manga. The metalanguage is intended to support a multimodal curriculum hence it is of importance that the metalanguage is not convoluted but accessible to students.

The theoretical approach informing this study is multimodal social semiotics. Since this theory views all semiotic resources as meaningful, it validates all texts and makes it acceptable to bring popular genres into the academic context. Moreover, the theory highlights the notion of meaning as designs. This is a key concept in this study as it supports the view of a metalanguage as a 'toolkit' which can be adapted and applied to other context of study. The idea of meaning as designs makes it viable to look at how a metalanguage for manga can be applied to the analysis of other visual narratives.

The study is also underpinned by a social theory of genre. This view of genre accounts for changes within genre conventions as well as the hybrid nature of genres. The theory

adequately explains the differences between Western comics and Japanese comics. It also adequately explains the nature of narratives.

The next chapter discusses the data and the method of analysis. This includes the proposed metalanguage for manga.

University of Cape Town

Chapter Three: Methodology

3.1 Overview of chapter

This chapter provides an overview of the research method and identifies the data used in the study. It also presents a framework for organising the data for analysis, namely Labov's (1972) narrative structure. The chapter ends with the method of analysis where the proposed metalanguage for manga is discussed in detail.

3.2 Overview of research method

Research is a process of systematic inquiry to enhance our understanding of the phenomenon being studied. The way in which we approach and carry out research is informed by our lifeworld and episteme. As Cohen et al point out “ontological assumptions give rise to epistemological assumptions; these, in turn, give rise to methodological considerations; and these in turn, give rise to issues of instrumentation and data collection (2007: 5). Ontology refers to our ‘being in the world’, how we experience and perceive ourselves in relation to our worlds and epistemology refers to our knowledge or philosophies of the world (Wisker 2008). Cope and Kalantzis (2006) refer to these terms as lifeworld and episteme respectively. If methods are derived from our theoretical assumptions of the world, it follows then that educational research is never neutral but “inextricably intertwined” with politics and decision-making (Cohen et al 2007: 5).

The epistemological assumption adopted in this study is constructionist. Constructionists view semiotic resources as “public” and “social” (Hall 1997: 25). They believe that meanings

are constructed through “symbolic practices” rather than simply existing (Hall 1997: 25). In other words, knowledge and meanings are seen as socially, culturally and historically constructed instead of a given reality. For this reason, representational resources are seen as pivotal to the construction and maintenance of our social realities. Since this study is concerned with notions of ‘design’, particularly the processes of designing and redesigning (NLG 2000), the epistemological orientation of this study aligns with that of the constructionist approach.

3.3 Data

The text chosen for analysis in this study is *Naruto*. This is a manga series written and illustrated by Masashi Kishimoto. *Naruto* is chosen as the data of analysis primarily because of its current popularity. The series is among best-sellers and in 2006, the English adaptation of the series won the ‘Quil Award’. This is a consumer driven award which aims to promote reading and literacy (Wikipedia 2009).

First published in 1999, *Naruto* is what Branigan would describe as a ‘simple narrative’. This is a narrative that is made up of “a series of episodes collected as a focused chain” (Branigan 1992: 19). The term ‘focused chain’ describes “a series of cause and effects with a continuing center” (Branigan 1992: 19). *Naruto* is named after the main protagonist of the story. It is therefore evident that this is a character-driven narrative with Naruto at the centre of each episode. To date, the episodes have been compiled into forty-six volumes or books. The manga series is still ongoing. The first episode of the first volume is used in this study.

It is important to mention that there are various English translations of *Naruto*. The officially translated version authorised for publication is distributed by *Viz Media*. In this edition, various aspects of the text have been changed to accommodate the target audience. In South Africa, this edition can be found in large book shops such as *Exclusive Books*, however they are expensive. The version that college students are more likely to read is the fan-translated edition. Fan-translated manga are unauthorised scans of the original manga. In most cases, they are freely available online. Besides the cost incentive, students who read manga are more likely to access the fan-translated versions as these are more up-to-date in comparison to commercial releases. Fan-translated manga is generally available soon after the latest releases in Japan. In contrast, commercial manga takes longer to release due to the translating and repackaging process. In addition, manga enthusiasts have a preference for the fan-translated version because much of the text is kept as close to the original as possible. Layout, names and sound effects, for example, are often unaltered in fan adaptations. Although students are more likely to read the fan translation of *Naruto*, this study makes use of the official English publication. The reason for this is that the official version accommodates readers who are unfamiliar with reading manga. Sound effects, names, layout are some aspects of the text that have been changed to suit readers who are unaccustomed to Japanese language and culture. However, where there are significant alterations to the text and these alterations greatly change the interpretation of the narrative, the differences will be discussed.

The aim of this analysis is to examine how various semiotic resources can be used to narrate a story, as well as to develop an approach to looking at other print-based visual narratives using the proposed metalanguage. The data analysis is structured using Labov's (1972) narrative structure.

3.4 Framing the Data: Labov's Narrative Structure

Story and plot are the basic elements of a narrative. They are intricately linked and should always be studied together. Labov's proposed narrative structure is one that ties together story and plot. Although based on oral narratives of personal experience, theorists have noted how the proposed structure applies to most narratives (van Leeuwen 2005; Branigan 1992). Labov's narrative structure is based on the assumption that the events that make up a narrative serve certain communicative functions. In van Leeuwen's words, the events serve "as something the storytelling *does* for the listener (or reader, or viewer)" (2005: 125). These communicative functions are as follows:

- a. Abstract: what was this about?
- b. Orientation: who, when, what, where?
- c. Complicating action: then what happened?
- d. Evaluation: so what?
- e. Result: what finally happened?

A sixth component, 'coda', is optional and may or may not be found at the end of narratives. Coda functions to "bridg[e] the gap between the moment of time at the end of the narrative proper and the present" (Labov 1972: 365). In other words, it functions to draw the audience from the place and time of the story to the place and time of the present where the narrative is being told.

3.5 Method of analysis

In social semiotics, the sign is seen as a social system of meanings. Meanings are made through semiotic resources that are grounded in their context of use. As mentioned before, systemic functional linguistics (SFL), as proposed by Halliday, is a social semiotic approach to the study of texts. This theory is grounded in the notion that texts are structured to perform certain functions in a given social context. SFL is a model of grammar that is structured to investigate “the organization of meaning according to the communicative functions that semiotic systems have evolved to fulfill” (Stenglin 2009: 36). According to Halliday, there are three communicative functions or ‘metafunctions’ of a text. These are ideational, interpersonal and textual. The ideational metafunction realises meanings through the elements that are represented. These include the represented participants and the context of situation in which they appear. The ideational metafunction is in fact a combination of two metafunctions: the experiential metafunction and the logical metafunction. The experiential metafunction describes meaning made through the kind of experience that is represented. The logical metafunction, on the other hand, describes the logical relations of the represented event. Next, the interpersonal metafunction deals with nature of the relationship between the producer and the receiver of the text. The textual metafunction looks at how meaning is distributed as a whole. Together, the three metafunctions are an attempt to account for the dimensions involved in the meaning-making process.

Halliday’s SFL metafunctional principle has been adapted by various theorists for the study of various semiotic resources besides language. Kress and van Leeuwen’s (1996) metalanguage for visual images is an adaptation of Halliday’s framework. Correlating with

Halliday's 'ideational', 'interpersonal' and 'textual' metafunctions, the metafunctions proposed by Kress and van Leeuwen are 'representational', 'interactive' and 'compositional'.

The proposed metalanguage has been devised by looking at manga texts in relation to Kress and van Leeuwen's framework and then adapting it accordingly. As mentioned before, Kress and van Leeuwen's framework is constructed for the analysis of single-framed, still visual images. Manga, however, makes meaning through sequential images. To account for this, this study draws on Matthiessen's (2007) work on 'rhetorical relations' in sequential images. This is a framework which looks at how sequences of images can be developed within a text. The framework aligns well with Halliday's logical metafunction and so it is fitting to foreground both the experiential and the logical metafunction within the representational metafunction. The metalanguage also draws a substantial amount of terminology from film because manga is a comic genre strongly influenced by it. Therefore, it would be practical to describe the images using film terms.

3.5.1 A Metalanguage for Manga

The Representational Metafunction

The representational metafunction is the level of a text which is concerned with what is represented. This includes the nature of the represented objects, participants and the context in which they are represented. The representational metafunction operates on two levels: experiential and logical. The experiential component of the metafunction is concerned with "the phenomena of the world as categories of experience" (Baldry and Thibault 2006: 22). In other words, it is concerned with how experience is represented. The logical component of the representational metafunction, on the other hand, is concerned with "the relations of

causal and temporal interdependency” (Baldry and Thibault 2006: 22). In other words, it is concerned with the logical progression of connected events.

According to Halliday, representations which produce particular meanings are always seen as “the expression of some kind of a process” (Halliday and Hasan 1985: 18). The experiential metafunction functions to interpret the kind of experience that is represented in an event. It is primarily concerned with how the world is categorised according to social experiences. Genres are products of a social process represented through text. This suggests that certain features of a genre can serve as indicators of social experience. In a narrative genre, the impression of a particular social experience is related through the story component of a narrative structure. Elements such as props, costume, location, colour and sound (in the form of onomatopoeia in written text) all contribute towards creating a representation of some kind of a social experience. This can be seen in Figure 1.

Figure 1 is a representation of a daily morning routine. The panels are read from left to right, top to bottom. In the first frame, props such as a bed and a pillow allow the reader to recognise the location as a bedroom. The beam of light shining through the window suggests that it is morning. The boy’s stretched hands, the yawn on his face, accompanied with the sound “yaaawn” indicate that the boy has just woken up. In the second row of frames, the milk and sandwich suggest that the meal being consumed is breakfast. From the first frame to the second last frame, in which the boy has changed his attire, there has been no disturbance. All these elements together create the notion of a daily routine.



Figure 1: Representing a social experience (Kishimoto 2007: 86)

Body posture, gesture and facial expression too function to relay a particular social experience. Look at Figure 2, for example. The feeling of anxiety and nervousness is established through the boy's hunched position, the sweat drops and the finger in the mouth.



Figure 2: Portraying social experiences through body posture, gesture and facial expression (Kishimoto 2007: 19)

While the experiential metafunction is concerned with the representation of a social experience, the logical metafunction is concerned with “the relationship between one process and another” (Halliday and Hasan 1985: 45). According to Baldry and Thibault, the logical metafunction is “realised by recursive structures which add one element to another so as to build up more complex chain-like structures” (Baldry and Thibault 2006: 22). This means that the logical metafunction is interested in how elements of a text link together to form a coherent whole. It can also be understood as how elements of a text push a narrative forward. The function of the logical metafunction aligns with Matthiessen’s (2007) notion of ‘rhetorical relations’. According to Matthiessen, rhetorical relations are concerned with “the development of sequences of passages in a text” (2007: 33). Drawing on Halliday’s (1994) theory of ‘projection and expansion’, Matthiessen discusses how through projection and expansion one image may develop another image.

‘Expansion’ refers to the augmentation of an image. There are three different levels in which an image can be augmented: elaboration, extension, transition. Elaboration refers to the restatement of an image. To elaborate something is to build onto something where the foundation is still the same. This can be understood as depicting the same image again but in greater detail or context. For example, an image may be elaborated through the means of

zooming in or out. That is, an image can be scaled down to a smaller magnitude where although the context of the image is minimised, the represented element is afforded greater detail. Alternatively, by zooming out, greater context is disclosed but there is less focus on the represented element.

While elaboration works with the same image but either develops it by closing in or moving out, extension propels the narrative forward by providing a new image. The new image ‘extends’ the existing image by providing additional information. Although the new image is ‘new’ in the sense that it has not been seen before, it is still related to the previous image. For instance, in film terms, extension can be realised through panning (camera moving sideways from a fixed position), tilting (camera moving up or down from a fixed position) or tracking (camera follows an object or participant). Figure 3 is an example of extension. The top image establishes the time of the day (night) but it does not provide details with regard to the setting. The bottom frame provides this information through lowering the camera eye. This is an example of extension through tilting.

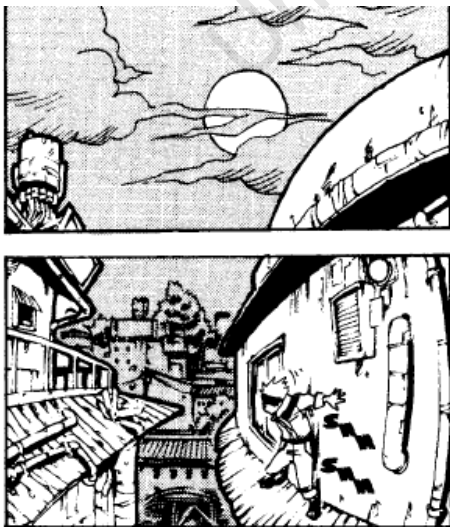


Figure 3: Extension through tilting an image (Kishimoto 2007: 24)

Transition is the augmentation of an image whereby there is a change in time and space. The word used by Matthiessen (2007) is ‘enhancement’. In this case, the narrative is carried forward by a change in time or space. In language-based narratives, this typically happens at the beginning of chapters or sections within a chapter. Since this process requires a change either in time or space, this study considers ‘transition’ a more fitting term. In images, the transition from one image to another can occur through techniques such as flashback or flashforward and split frames which evoke the notion of ‘meanwhile’. Any image in which a temporal or spatial change has occurred in relation to a previous image can be qualified as a transition. Figure 4 is an example of where a narrative is propelled forward through transition. Reading from right to left, the narrative was initially situated at a traditional Japanese restaurant in the evening. The narrator exits this scene in the narrative with an external view of the restaurant. It is appropriate that this view should be from that of an omniscient narrator. In the text frame, puffs of cloud indicate that there is a shift in time. It is no longer the evening but day time. The last frame takes the reader into a classroom. These three frames clearly indicate a shift in time and space.

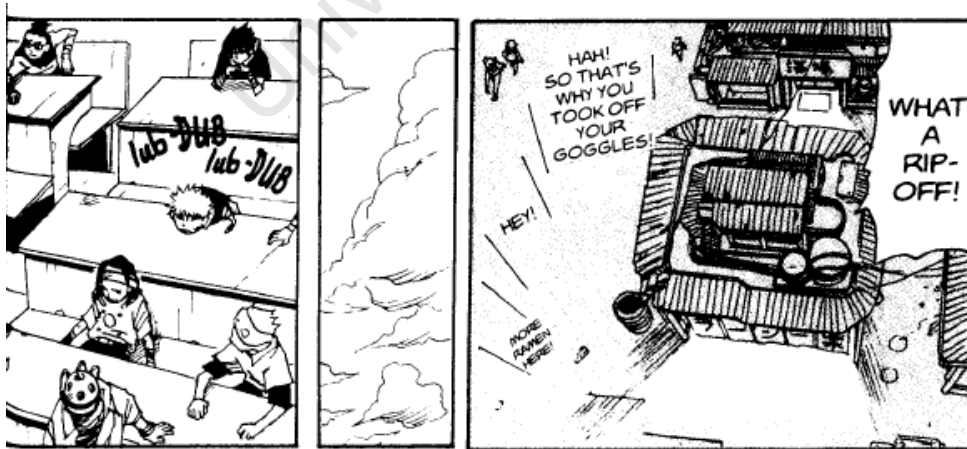


Figure 4: A shift in time and transition through transition (Kishimoto 2007: 18)

Another way of pushing the narrative forward is through dialogue, whether it is internal dialogue (thought) or external dialogue (speech). This process is referred to as ‘projection’. Projection in comics is usually captured in frames. Figure 5 is an example of projection through voiceover narration.



Figure 5: Projection through a voiceover (Kishimoto 2007: 17)

Figure 5 depicts two people having a discussion. In Frame 1, the ninja on the left asks the boy a question. The boy proceeds to answer this question in Frame 2. He goes on to elaborate his answer and in doing so, his explanation is carried into frame 3. In Frame 3, the boy’s dialogue is overlapped with an image that is set in a different time and place. The image would have been otherwise out of context in this sequence of images but the boy’s dialogue serves as a continuity element, propelling the narrative forward.

A framework summarising the representational metafunction is shown in Table 1.

<p><u>Representational Metafunction</u></p> <p>Experiential</p> <p>Costume Props Location Colour Sound Effect (onomatopoeia) Body posture, gesture and facial expression</p> <p>Logical</p> <p><i>Expansion</i></p> <p>Elaboration (same image but in greater detail or context)</p> <ul style="list-style-type: none">• Zoom in /out <p>Extension (provide new but related information)</p> <ul style="list-style-type: none">• Pan• Tilt <p>Transition (change in time and space)</p> <ul style="list-style-type: none">• Flashback / forward• Split frames <p><i>Projection</i> (articulation of speech or thought)</p> <ul style="list-style-type: none">• Speech• Thought

Table 1: The representational metafunction

The Interactive Metafunction

The interactive metafunction is concerned with the interpersonal relationship between the audience and the participants of the represented world. Specifically, it is primarily concerned with the social relations and the evaluative orientations between the two interactants. Kress and van Leeuwen (1996) identify three particular categories which express the interactive metafunction: contact, social distance, attitude. This study has added facial expression, body posture and gesture as another category in this framework.

Contact is concerned with the presence or absence of a gaze from the represented participant. In film terms, this can be regarded as viewer identification and it is established through camera positioning. The camera can be placed in a number of positions depending on who the director wishes the viewer to identify with. This camera positioning is also known as 'point of view'. This study replaces Kress and van Leeuwen's (1996) term, 'contact' for 'point of view' since the latter is more commonly used. A first person point of view is established when the camera takes the position of the represented participant. This position allows the viewer maximum identification with the represented participant. Figure 6 is an example of a subjective point of view where the depicted images simulate the vision of an eye opening. The image begins with a totally black frame which imitates the vision of the closed eye. A horizontal central split in the next image represents the eye partially opening. This 'eye' widens until the entire image is clear in the last frame. In this representation, there is a strong interpersonal engagement with the reader as the reader adopts the position of the character and views the world at the exact moment as the character.

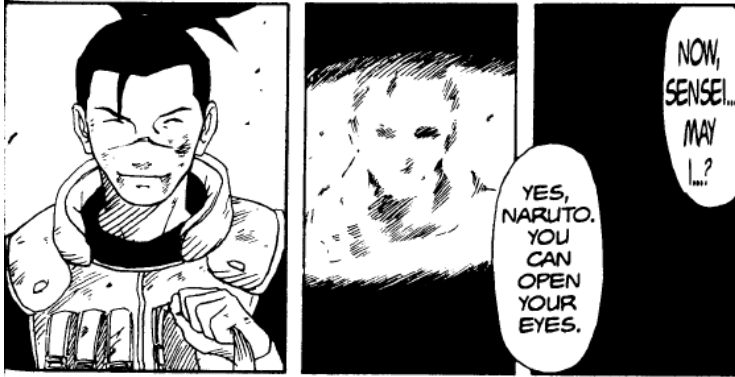


Figure 6: A first person subjective point of view (Kishimoto 2007: 58)

Third person point of view is established when the camera appears to view the narrative unfold from a third person in the narrative. Viewer identification with the represented participant is less strong from a third person perspective. The omniscient point of view presents a narrative unfolding from ‘god’s’ view or a bird’s eye view. This is a point of view that does not belong to anyone in the narrative. Viewer identification is absent from this viewpoint. For example, Figure 7 and Figure 8 are images of the same location. In Figure 7, the perspective is from that of a third person. It is as if someone is looking at the restaurant from the outside. In Figure 8, the restaurant is viewed from above, from a bird’s eye view. This is not a position anyone in the narrative can occupy, thus it can be said that the point of view comes from that of the omniscient narrator.



Figure 7: A third person point of view (Kishimoto 2007: 16)

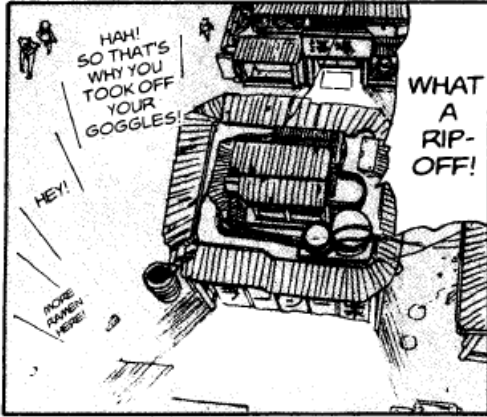


Figure 8: An omniscient point of view (Kishimoto 2007: 18)

An image is not only always seen from a particular point of view but also from a particular distance. The term 'social distance' refers to the proximity between the subject depicted and the audience. Kress and van Leeuwen (1996) identify three types of proximity: intimate/personal, social, impersonal. In film, social distance is realised through the type of shot used: an intimate/personal distance through a close up shot, a social distance through a medium shot and an impersonal distance through a long shot. The type of shot used, however, not only reveals the social distance, but it can also establish how much information is disclosed in an image. For example, a long distance shot will typically reveal more of a character or environment than a close up shot. Therefore, while a long distance shot may reveal less of a character, and place a distance between the character and the reader, it will nonetheless provide the reader more information about the character's environment. In film terms, a long shot which establishes the scene is referred to as the establishing shot.

Another resource that is closely related to point of view and distance is the angle from which an image is depicted. The term 'attitude' (Kress and van Leeuwen 1996) refers to the interpersonal attitude, specifically the power and the level of involvement, realised through

the positioning of camera angles. Relationships of power are constructed through the use of vertical angles (low, medium, high). For example, a low angle depicting the represented participant looking down at the viewer can convey power over the viewer. In contrast, a high angle which depicts the viewer gazing down at the represented participant can afford the viewer power. An angle which is placed at eye level can establish a feeling of equality.

The level of involvement which the viewer has with the world of the represented can be constructed through the use of horizontal angles (frontal or oblique). Frontal shots create maximum engagement as the viewer is directly confronted with the world of the represented. In contrast, oblique shots suggest detachment as the viewer looks at the world of the represented from the side. It is important to note that these are not necessarily the exact meanings of these angles. Rather, as Jewitt and Oyama write, “[t]hey are an attempt to describe a meaning potential, a field of possible meanings, which need to be activated by the producers and viewers of images” (2001: 135). The potential meanings described here are derived from frequent usage in Western images, yet these angles have the potential to mean otherwise in particular contexts.

Facial expression, body posture and gesture are other resources which express the interactive metafunction. The human form is one of the most emotionally expressive resources that invite the reader to take on particular evaluative stances. This is especially effective when the interaction between the represented element and the reader is in the form of direct address. Figure 9 is an example of point of view working with gesture, inviting the reader to engage directly in the narrative. The subjective point of view along with the index finger pointing straight at the reader invite the reader into the narrative world. The resources demand a

first-person interaction. A framework summarising the interactive metafunction is shown in Table 2.



Figure 9: The index finger demands interaction from the viewer (Kishimoto 2007: 12)

Interactive Metafunction

Point of View

- Objective – omniscient (a point view of that does not come from anyone in the narrative)
- Point of View – 3rd person
- Subjective – 1st person (from the character's perspective)

Social Distance

- Intimate – Close Up
- Social – Medium
- Impersonal – Long Shot

Attitude

- Vertical angles: Low angle – maximum power
High angle – lack of power
Eye level – equal power
- Horizontal angles: Frontal shot – engagement
Oblique shot – detached

Facial expression, body posture and gesture

Table 2: The interactive metafunction

Compositional Metafunction

The compositional metafunction is concerned with the composition of the text in its entirety. It is concerned with “the way in which the representational and interactive elements are made to relate to each other, the ways they are integrated into a meaningful whole” (Kress and van Leeuwen 1996: 181). A key concept which affects the compositional meaning is layout. Kress and van Leeuwen identify three inter-related systems which are all concerned with the layout of a text: information value, salience and framing.

Information value describes the value attributed to the different positions of a page. For example, the centre is generally regarded as having more value than the margin because we tend to focus more on elements in the central area. Certain meanings are also attributed to the horizontal (left/right) and vertical polarisation (top/bottom) of a page. Kress and van Leeuwen (1996) suggest that conventionally, the Given, something which is familiar, is usually placed on the left side of a page. In contrast, the New or the unfamiliar is placed on the right. The top is usually regarded as the Ideal while the bottom is regarded as the Real. The meanings of these polarisations are very much based on the reading practices of the West. This differs greatly to the reading practices of the East where manga, for example, is read from right to left, top to bottom. Hence, it may be that the meanings and the values attributed to the positions of a page may differ in manga. Nevertheless, according to Kress and van Leeuwen,

All cultures work with margin and centre, left and right, top and bottom, even if they do not all accord the same meanings and values to these spatial dimensions. And the way they use them in their signifying systems will have relations of homology with other cultural systems, whether religions, philosophical or practical (1996: 199).

This suggests that while cultures with a different reading direction to the West may attribute different values to the various positions, certain values could also overlap as a result of a number of factors. These days, especially, practices are increasingly merging as a result of a globalised context.

The notion of 'saliency' applies to the layout of a text where the represented elements are given visual prominence through particular techniques. For example, foregrounding and backgrounding can be achieved through the allocation of space. Elements at the forefront tend to be larger in size than those in the background. The larger the size of the represented element, the more prominent it is. Saliency can also be achieved through colour saturation. Another method of achieving saliency is through focus. That is, visual saliency can be achieved by blurring or sharpening the focus of a represented object or participant. In sum, saliency applies to the "degree to which an element draws attention to itself, due to its size, its place in the foreground or its overlapping of other elements, its colour, its tonal values, its sharpness or definition and other features" (Kress and van Leeuwen 1996: 225). In fact, Centre/Margin positioning also applies to the notion of saliency. Elements placed at the centre of a page have more value because they draw more attention than elements placed in the margin.

Framing describes the use of elements such as border lines. Frames can function to link or detach elements of a text, "signifying that they belong or do not belong together in some sense" through its presence or absence (Kress and van Leeuwen 1996: 225). Framing is an essential compositional device for manga and comic art in general. Sequential frames create a flow in the narrative by segmenting one moment followed by another. This creates the illusion of the passage of time and the notion of cause and effect. As Ryan writes,

The reader (for the eye movement amounts to an act of reading) constructs a story line by assuming that similar shapes on different frames represent common referents (objects, characters, or setting); by interpreting spatial relations as temporal sequence (adjacent frames represent subsequent moments); and by inferring causal relations between the states depicted in the frames (2004: 141).

The actual shape of the frames also has meaning-making potentials. According to Baldry and Thibault, frames function as a “metacomment” which helps to signify how the world depicted in the frame should be interpreted (2005: 10). For example, in Figure 10, the character is framed at a distorted angle. The tilted view reflects the character’s distorted world. In Figure 11 the diagonal frames simulate the action happening at that moment. It creates the illusion of daggers slashing through the page. Even the frames of speech bubbles can make a metacomment about the dialogue framed within it. In Figure 12, the jagged framing reflects the anger, intensity and volume of the voice.



Figure 10: A canted angle signifies a world that is distorted (Kishimoto 2007: 35)



Figure 11: The diagonal frames simulates the action (Kishimoto 2007: 31)



Figure 12: Jagged speech frames signify the intensity and volume of voice (Kishimoto 2007: 14).

This study includes typography as another aspect which affects the composition of a text. In the past, it was generally regarded as “a transmitter of the written word” (van Leeuwen 2004: 14). Today, however, it is evident that typography transmits more than the written word. According to van Leeuwen, typography is a “communicative mode in its own right” (2004: 14).

It no longer communicates only through variations in the distinctive features that allow us to identify and connect the letterforms, and not even only through the

connotations of particular fonts, for example, the association of Park Avenue script with formality and high status, but also through modes which it shares with other types of visual communication – color, texture, and movement (van Leeuwen 2004: 14).

This is particularly evident in comic art where the written text is often treated as an extension of the imagery. Drawing from Lim (2004), Unsworth refers to this as “homospatiality” (2006: 61). According to Unsworth, homospatiality “refers to texts where two different semiotic modes co-occur in one spatially bonded homogenous entity” (2006: 61). For example, in Figure 13, the words ‘drip drip’ function both as sound effect and visual effect. The sound of the blood dripping is conveyed in the linguistic representation. The typography visually reflects the blood dripping.



Figure 13: “Homospatiality” (Kishimoto 2007: 50)

In fan-translated manga, sound effects are mostly left untranslated. It is interesting to note how the reader conjures up the sound from the context of the situation and from the choice of font. Although the meaning is subjective to the individual reader, he/she is encouraged to associate it with particular sounds from the typography. For example, examine the sound

effects in Figure 14. In example A, the puffs of cloud besides the written text suggests that this sound is along the lines of “poof”, either the sound of the pipe or in reading the frame in context, the sound of a hat being placed on the head. In example B, the squiggles evoke the sense of something ‘girly’, so the sound is something probably along the lines of “mwah”, the sound of a kiss. In example C, the jagged edges of the letters and the thick bold font express force. The sound is probably along the lines of “kkkkkkk”, the sound of shoes sliding through gravel. Compare these to the English edition in Figure 15. In the English version, the typography appears slightly different. This could be due to the difference between the Japanese writing system and the English alphabet. The typography has possibly been changed to suit the different lettering systems.

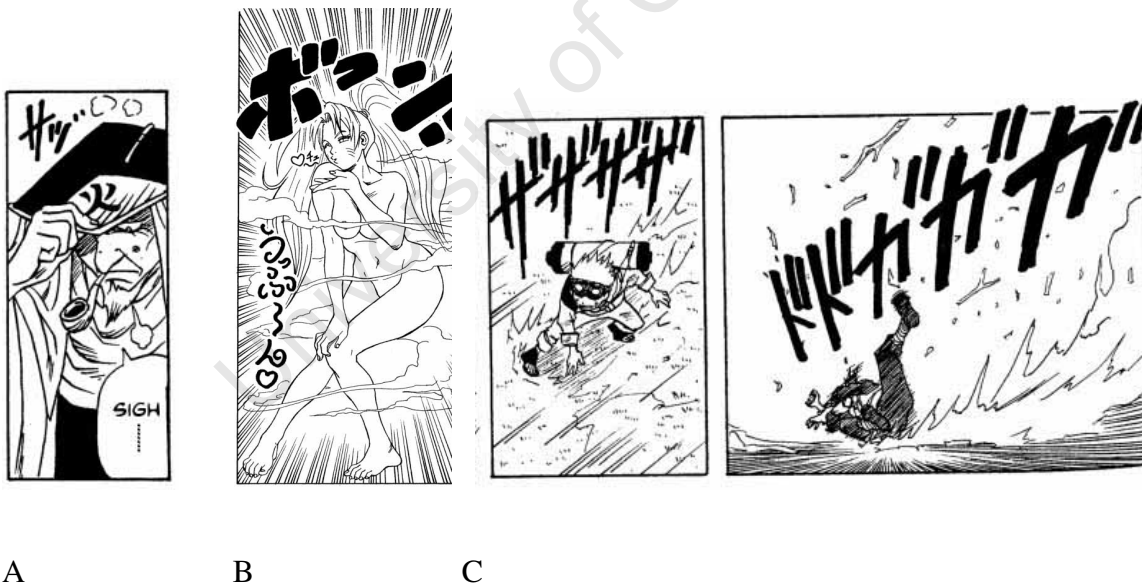


Figure 14: Typography as an important resource to conjure up the sound effect (Kishimoto 1999: 9, 14, 45)

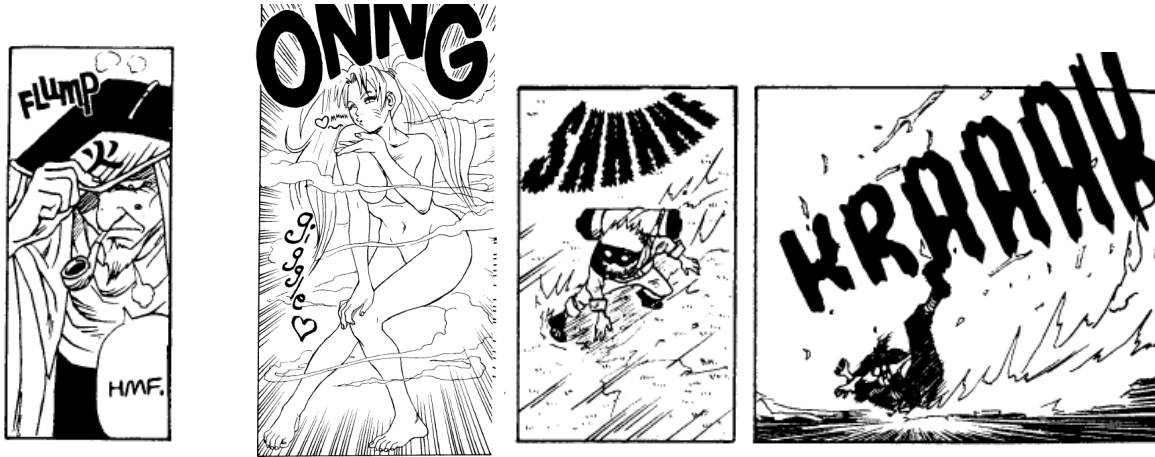


Figure 15: English translations of the Japanese sound effects (Kishimoto 2007: 9, 14, 45)

The compositional metafunction is summarised in Table 3.

Compositional Metafunction

Information Value

- Centre/Margin
- Ideal/Real
- Given/New

Salience

- Space proportion
- Colour Saturation
- Focus

Frames

Typography

Table 3. The compositional metafunction

This chapter has presented an overview of the research methodology including the proposed metalanguage for manga. Using the proposed metalanguage, the next chapter will analyse the data of the study and address the research question “how are semiotic resources used to narrate a story in manga?”

Chapter Four: *Naruto* from a social semiotic perspective

4.1 Introduction

This chapter takes a social semiotic approach to analysing the first episode of *Naruto*, the manga. As mentioned in the previous chapter, various aspects of the text have been altered in the English edition of *Naruto* to accommodate the Western audience. One of these changes is the layout of the text and this is particularly evident in the abstract. The change in the layout has a considerable impact on the reader's experience of the narrative. In this section, both the Japanese and English editions of the abstract are discussed. The analysis begins with the Japanese edition, the version as told and depicted by the author of the manga. To overcome the language barrier, the study draws on the fan-translated version.

The aim of this data analysis is to investigate how semiotic resources are employed to narrate a story. Thus while a discussion of the semiotic resources will be foregrounded, the study will also attempt to read the data as if reading a narrative. The narrative has been divided into five sections. The sections correspond with Labov's five events of a narrative structure: abstract, orientation, complication action, evaluation and resolution. It is important to note that in manga, sequential images are read from left to right and top to bottom.

4.2 Abstract

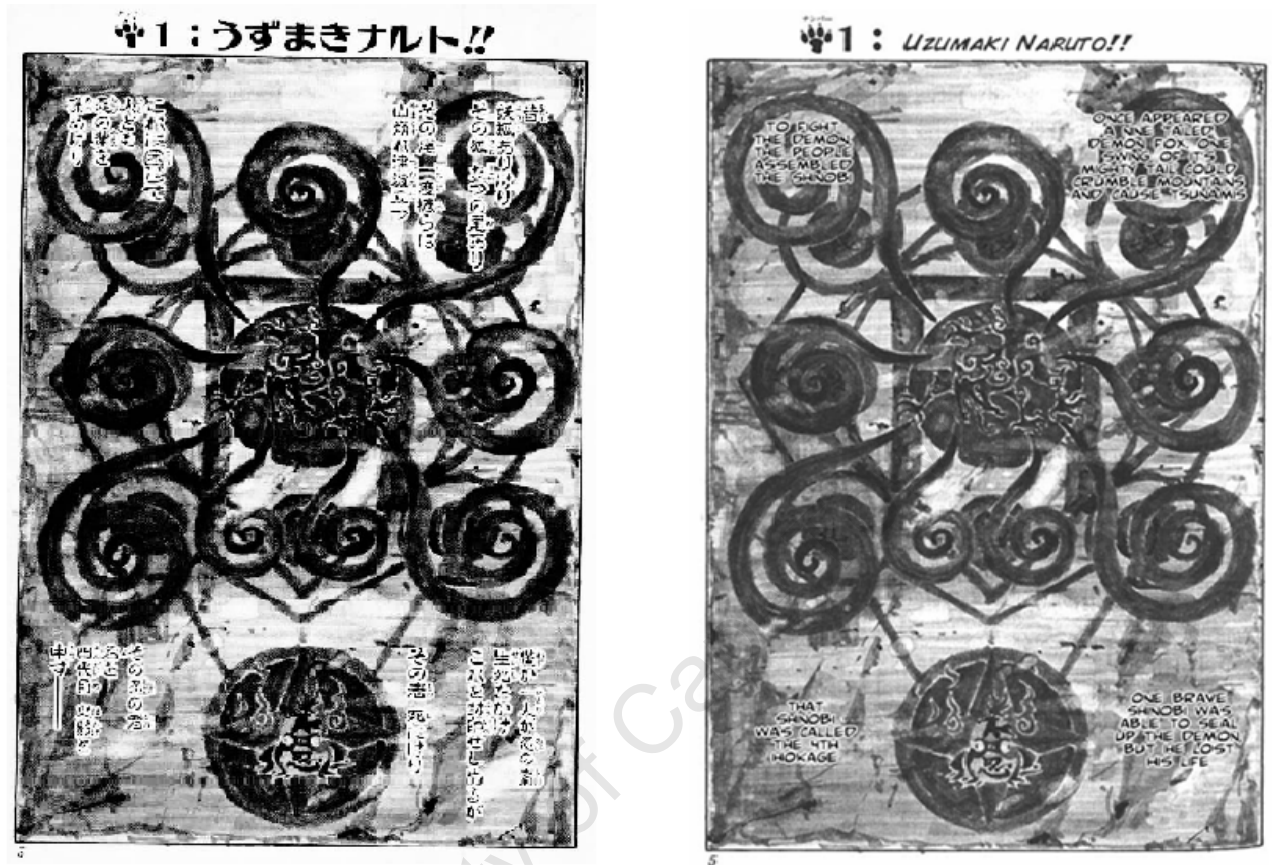


Figure 16: The original *Naruto* and the fan English translated version (Kishimoto 1999: 4)

The abstract announces the opening of a narrative and has the function of summarising the story that is to follow. Figure 16 is the abstract to *Naruto*. Here, the summary function is carried out by the chapter title “Uzumaki Naurto!!” Uzumaki Naurto is the name of the main protagonist in the story. The chapter title therefore clearly indicates that the first episode is centred on the main character. In language-based narratives, chapter titles are usually separated from the actual body of the narrative through white space. In this case, it is separated from the body by a frame. The heading of this chapter title is more decorative than most. The paw icon and the double exclamation mark are extra features and they function as visual representations of Naruto’s character. The paw icon serves as a visual symbol for

Uzumaki Naruto as he literally has a demon fox sealed within him. Naruto's hyper-active personality is reflected in the double exclamation mark which suggests extreme emotion and energy. In this manner, the chapter title presents Naruto as the focus of the story as well as provides a summary of his character.

The chapter title is distinctively separated from the body of the text by a frame. This frame captures the world of the narrative within the frames and clearly makes a distinction between the world outside the narrative and the world inside it. According to Baldry and Thibault, frames provide a "metacomment on the depicted world of the picture" and specify "how the things inside the frame are to be taken" (2006: 10). In this case, the frame visibly separates the world of the narrative from the world outside and establishes a boundary between the real and the fiction. The frame thus indicates that the reader is crossing into the world of fiction so this world is not to be taken for real.

Inside the frame, the Centre/Margin composition of the text draws the reader's attention to the various shapes, swirls, lines and symbols at the centre of the page. This image evokes a sense of an ancient and mysterious world. There is a pattern to the abstract image. The pattern conveys a sacred code as the orderliness suggests that they are placed in those positions by design rather than by chance. Projected through the voice of an omniscient narrator (implied because there are no frames binding the words to a specific speaker), the written text unlocks some of this code.

According to the omniscient narrator, there was once a demon fox who had nine tails. Ninjas (Shinobi) were called upon to overpower this destructive demon fox. In the end, the fox was captured and sealed but the ninja who effected this lost his life. Reading the omniscient

narration in relation to the image, the symbols come to take on particular meanings. There are exactly nine swirls with their tails all directed towards the centre circle. This suggests that the symbol at the centre is of importance. The nine swirls can be said to represent the nine tails of the demon fox and the central circle, the demon fox. The latter is boxed in by a thick frame and this thick frame is boxed in by other frames. Two of the lines that function as the outermost frames are connected to another circle at the bottom of the page. This circular form appears to hold the frames in place. At the centre of this circle is a Chinese word meaning tolerance. This is a symbol commonly associated with ninjas. The image can be verbally translated as the seal binding the demon fox in its container. Although the written text is essential in unlocking the meaning of this image, its role here is to support the image. This is reflected in the composition of the text.

On the compositional level, the image is given preference as a result of its size in comparison to the written text and importantly, its position at the centre of the page. Kress and van Leeuwen (1996) point out that while central composition is less common in the West, this organisational principle is employed frequently in the East. They attribute this to “the greater emphasis on hierarchy, harmony and continuity in Confucian thinking” (Kress and van Leeuwen 1996: 206). Certainly, this ideology is reflected in the narrative. The central positioning of the image establishes the mood of the narrative at the outset. The swirls, lines and symbols evoke the sense of an ancient culture, a level of fantasy and mystery. The swirls, lines and symbols take on specific meanings as they are read in conjunction with the written text. Spread out on four corners as if to anchor the page, the written mode pins down the meaning to the otherwise abstract image. The role of the image is to arouse the reader’s curiosity and the written text is meant to satisfy it.

Tempo is another element that is evoked through the layout of the written text. By placing the written text on all four corners of the page, it requires the reader to follow the narration from corner to corner. In doing so, it creates pauses to the narration and establishes tempo. In this case, the rhythm in the narrative is created through the use of space. In the English edition (see Figure 17), the pace of the narrative is established through framing. Frames connect or disconnect elements of a text (Kress and van Leeuwen 1996). Thus by boxing sentences in frames, it disrupts the flow in the narration and establishes visual pauses.

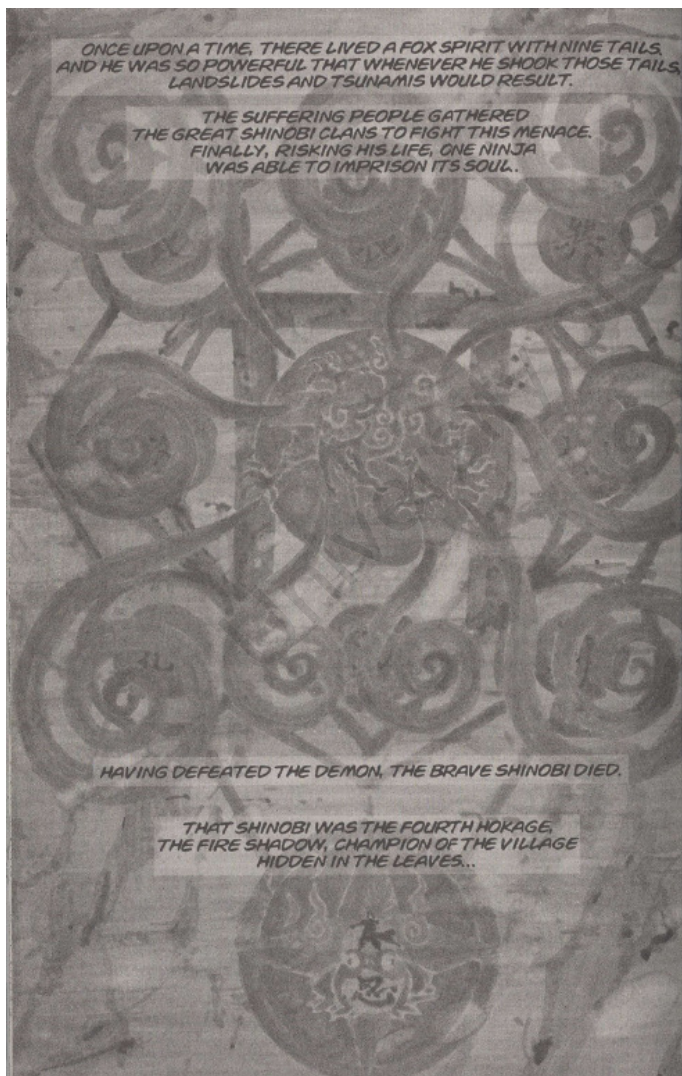


Figure 17: Abstract from the English edition (Kishimoto 2007: 4)

In the English edition of *Naruto* (Figure 17), the layout has been altered and this results in a very different reading experience. The chapter title with the summary of the narrative is omitted here. Instead, the abstract leaps straight into the narrative world. A possible reason for this is that the title would be seen as redundant since it is repeated on the next page. Cultural practices play an important role in the composition of a text (Jewitt and Oyama 2001). This text is situated in a Western context, and it is perhaps more of a Western practice to discourage redundancy. Kress and van Leeuwen (1996) mention that in contrast to the Centre/Margin composition of texts in the East, Western text compositions tend to polarise elements. This is certainly the case in this instance. In this version of the abstract, the written text is positioned at the top and bottom of the page. This positioning encourages the reader to read from top to bottom. Since this is the customary reading path in the West, the reading direction may have been altered to suit the Western reader. In changing the reading path, the value previously attributed to the image is altered. The image is no longer the central element despite the fact that it is still positioned at the centre of the page.

In contrast to the Japanese edition, the written text is established as the key element in this abstract. A number of compositional techniques have been used to guide the reader to the written text first. For one, the written text is boxed in frames and superimposed over the visual image. This contrasts with the Japanese edition where the writing is framed by the pattern. The visual image is also printed in a lighter shade compared to the written text. It appears as if its function is to serve as 'wallpaper', a decorative element to the narrative. This notion is reinforced by the fact that the written text is in a bold font and runs across the pattern.

By directing the reader's attention to the written text first, he/she begins the narrative with facts. The reader is told there was once a destructive fox spirit who caused great suffering to the people. Ninjas were called upon to subdue this fox and one ninja was eventually able to imprison its soul. The idea of imprisonment is visually supported by an image at the centre. After this visual break in the narration, the narrator continues to inform the reader of the ninja's identity and how he died. A visual symbol for ninja follows the verbal narration. The composition of this abstract places the written text as the principle element. The image plays a supportive function as it visually elaborates on the written text. This not only removes the mystery from the visual image but it establishes a strong sense of hierarchy. The written text is presented as the facts of the story, and in doing so it affords the narrator a strong authorial voice. In contrast, the narrator's voice is less authoritative in the Japanese edition as a result of its position at the corners of the page. By foregrounding the image, the abstract arouses the reader's curiosity. In comparing these two abstracts, it emerges that textual composition is socially situated and it plays an important role in establishing the mood and the reading experience.

4.3 Orientation

Narratives are different from other types of genres because they are characterised by time and by cause and effects. As discussed earlier, a narrative is composed of both story and plot. 'Story' is the chronological order of events as it happened while 'plot' is the chronological order of events as it is told. A 'story' is converted into a narrative as the plot purposefully structures the events into a meaningful order. It is therefore expected that the concept of 'rhetorical relations' would be key in the study of narratives.

Rhetorical relations provide an explanation for text coherence. They are concerned with “how texts are developed” (Matthiessen 2007: 33). The theory that Matthiessen (2007) proposes for looking at the rhetorical relations in sequential images is based on Halliday’s logico-semantic relation types (projection and expansion). These relations are explored in detail in the following sequence of images.



Frame 1

Figure 18: Wide shot establishing the location (Kishimoto 2007: 9)

An orientation functions to set the scene of the narrative. This is done by specifying “the time, place, persons, and their activity or the situation” (Labov 1972: 364). Naruto begins with a wide shot which broadly introduces the reader to the larger setting of the narrative (see Figure 18). The story takes place in a village. The electricity wires linked from pole to pole suggest that it is set in modern times, yet the buildings ascribe to the ancient architecture of

the East. This sense of the Orient is supported by the Chinese character situated at the centre of the page. The character means 'fire' and it is a symbol that readers will come to recognise as having importance as the narrative progresses.

The reader's attention is guided to the facial engravings and the building with the fire symbol by a number of compositional resources. Firstly, these elements are positioned at the centre of the page. This central position is emphasised by the shading on the margins. The different shades of black and grey provide depth to the image as well as direct the eye to the centre due to the colour contrast. The colour also provides a sense of realism as it evokes the concept of 'light' and 'shadow'. A 'vector' is a line which connects one represented participant to another (Kress and van Leeuwen 1996). It "has properties such as *dynamic force*, *directionality* and *orientation*" (Baldry and Thibault 2006: 35). In this case, a vector is created by the roof ridges of the building on the right. This horizontal line guides the reader's attention to the building at the centre.

While the compositional resources in Frame 1 guide the reader's attention to the centre of the page, it is the representational resources at the centre which arouses the reader's curiosity. Even from a distance, it is evident that something is amiss with the facial engravings. The faces appear to have blood pouring out of their eyes and nose, there are spirals on the cheeks and the words 'IDIOTS' and 'FOOLS' are visibly scribbled across an engraving. This peculiarity cues the reader for some form of explanation and the next frame (Figure 19) provides the answer through an 'elaboration' of the image. In elaboration,

[one image] elaborates on the meaning of another by further specifying or describing it. The secondary [image] does not introduce a new element into the picture but rather provides a further characterization of one that is already there, restating it, clarifying it, refining it, or adding a descriptive attribute or comment (Halliday 1985: 203).

By magnifying the image so that the represented elements are seen in greater detail, it becomes evident that these marks are the result of vandalism. The culprit, a boy, is still laughing deviously at the scene of the crime.

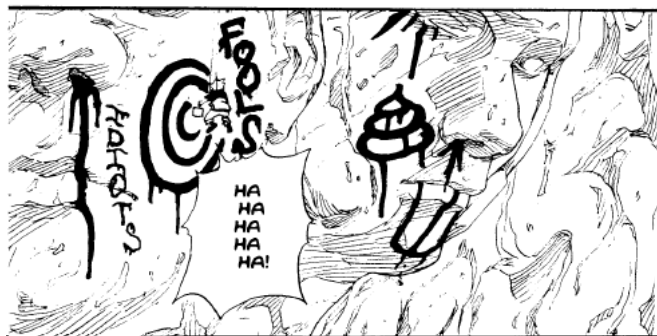


Figure 19: Specifying an image through elaboration (Kishimoto 2007: 9)

Both Frames 1 and 2 employ a framing technique commonly referred to as ‘bleeds’. In this technique, the frame extends beyond the boundary of the page and in doing so, creates the illusion of time standing still. Bleeds establish “the mood or a sense of place for whole scenes through their lingering timeless presence” (McCloud 1994: 103). This feeling of timelessness is felt more strongly in Frame 1 than in Frame 2 because of the space proportion afforded to Frame 1 (see Figure 20). In contrast, the presence of time is vividly felt on the next page (see Figure 21) where smaller and tighter frames evoke a strict sense of time. The next page introduces the reader to the characters and their situation.



Figure 20: Orientation, establishing the setting (Kishimoto 2007: 9)

As the reader turns the page, the narrative shifts to another scene where the characters in the story are introduced. Transitions expand an image through a temporal or spatial shift. McCloud (1994) refers to this type of image development as ‘scene-to-scene’ transition. This type of transition creates a considerable lapse in time and space and consequently weakens the flow of the narrative (Lim 2007). In this case, the transition occurs at the same time as the turning of the page. When a page is turned, there is also a lapse in time and the reader’s attention is likely to escape the narrative momentarily. It is therefore appropriate that the scene change coincides with the page turning. By incorporating this aspect into the composition of the text, it turns the act into a transitional element. The turning of the page thus becomes part of the reading experience.



Figure 21: Orientation, establishing the characters and their situation (Kishimoto 2007: 10-11)

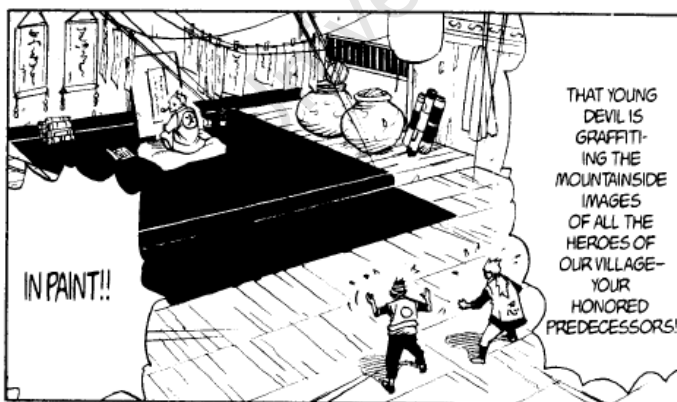
The reader enters the new scene with someone calling out ‘Lord Hokage!!!’ (see Figure 22). The triple exclamation mark and the jagged speech frame suggest that the tone is loud and urgent. Upon hearing this, the old man who is practicing calligraphy tenses up and a roll of sweat rolls down his forehead. His verbal response suggests that the address is directed at him. He is Lord Hokage. Despite his tense expression, he replies in a calm voice. This is suggested by the smooth frame of his speech bubble.



Frame 3

Figure 22: Close-up shot establishes a sense of intimacy (Kishimoto 2007:10)

Frame 3 is the first instance in the narrative where the reader is placed within an intimate distance from a character in the narrative. It invites the reader to engage in the narrative not only because of the intimacy afforded by the camera distance but also by the fact that the reader is positioned on the same level as the character. Nevertheless, the frame is too close to the action. It does not provide perspective regarding the relocation of the setting. Where is Lord Hokage situated and who is calling out to him? These questions are answered through an elaboration of the image in Frame 4 (Figure 23).



Frame 4

Figure 23: Re-establishing the setting through an omniscient point of view (Kishimoto 2007: 10)

Elaboration develops an image by describing it in greater detail. This can be done by either closing in on the image and magnifying the represented elements or ‘moving out’ of an image so a greater context is provided (Matthiessen 2007). Frame 4 (Figure 23) provides more detail with regard to the setting by retracting from the intimate position in Frame 3 and providing a point of view that is external to the narrative. The point of view is now very distant and positioned at an angle above the characters. This perspective is likely to be that of an omniscient viewer rather than a character in the narrative.

In Frame 4, the reader realises that the narrative has shifted indoors. Lord Hokage is seated on a stage which suggests that he holds a position of power. Two ninjas verbally elaborate on the event depicted in Frame 2. According to the ninjas, Naruto is defacing the monument of Lord Hokage’s predecessors. These people are heroes of the village so they are greatly honoured. Naruto, however, shows disrespect by his act of vandalism and the ninjas’ hysteria at this outrage is reflected in their body posture. The ninjas’ fury can also be sensed by the jagged speech frames and the vector lines drawn around them. These lines appear to emit from the ninjas and they reflect the intensity and the volume of their voices. The lines are used quite often in the narrative thus this study will refer to them as ‘volume lines’ hereafter. In Frame 5 (Figure 23), looking defeated, Lord Hokage sighs, puts on his hat and heads off to sort out the mess. The word “flump” written above Lord Hokage implies the sound of a hat being placed on the head.



Frame 5

Figure 24: A medium shot (Kishimoto 2007: 10)

It is interesting to note that when the two ninjas report to Lord Hokage of the event in Frame 2, they describe Naruto's actions as 'graffiting'. According to Kress (2003), the written mode and the visual mode demand different epistemological comments. The written mode demands a commitment to naming a relation while the visual mode demands a commitment to a location of space. The event in Frame 2 is described as 'graffiting' but what does 'graffiting' entail? Drawing? What kind of drawing? The word is vague and waits to be filled with meaning. The image, on the other hand, has to make a commitment and fill the word 'graffiti' with meaning. In this case, the act of graffiting entails painting the words 'idiots', 'fools', drawing spirals and other visual symbols on particular spaces on the monument. Thus, Kress comments that "images are plain full with meaning, whereas words wait to be filled" (2003: 4).

The move from Frame 4 to Frame 5 can be described as 'extension'. Extension expands an image by adding new but related information (Matthiessen 2007). The reader has already been introduced to Lord Hokage in Frames 3 and 4. Nevertheless, in both frames, he is presented either too close or too far. The close up and the extreme long shot of Frame 3 and 4 respectively do not provide a complete image of Lord Hokage. In Frame 5, a medium shot affords the reader an image of him from the waist upwards. This distance, which Kress and van Leeuwen describe as the distance that "subjects of personal interests and involvements are discussed" (1996: 130), presents a partial view of Lord Hokage's figure and at the same time keeps the reader at an intimate distance. This image adds to the narrative on an experiential level as the costume, an oriental style robe with the Chinese symbol, vividly evokes the sense of being situated in an ancient Eastern culture.

The communicative purpose of this second scene is to further extend the setting of the narrative. The atmosphere of the setting is emphasised by the props and costumes employed. The Japanese writing, calligraphy paint brush, scrolls, robe and hat, all evokes a sense of a traditional Japanese culture.

The second scene also functions to introduce the reader to the characters of the narrative. The boy who appeared briefly in Frame 2 is no longer just any boy but the protagonist, Naruto. He is evidently a trouble-maker. Lord Hokage is another character introduced. He is presented as an important person in the village. This is implied in his title 'Lord' and by his position, seated on the stage. It is no coincidence that the symbol for 'fire' reappears on his hat. Later in the narrative, it is revealed that the narrative world is divided up into five main countries. The countries are named after the five basic elements – fire, wind, earth, water and lightning. Each country has a hidden village of ninjas that safeguards it. The Fire country is protected by a village of ninjas known as the Hidden Village of Leaf. The head ninja of each hidden village is given the title 'kage'. The term 'ho' means fire so the 'hokage' is the leading ninja in the country of Fire. It follows that the fire symbol symbolises the country and only the Hokage can bear the symbol.

The narrative returns back to the main event in Frame 6 (Figure 25). Naruto is still painting on the mountain face and a crowd has gathered below him. They yell threats at him and tell him to stop, but he continues to paint. The reader can 'hear' his action by way of the words "swish swish". Social semioticians stress that signs are motivated (Halliday and Hasan 1985; Kress et al 2001; Kress 2000, 2003). Signifiers are chosen because of their aptness in expressing particular meanings rather than for arbitrary reasons (Kress et al 2001; Kress 2000, 2003). The words 'splash splash' could just well have been used here to imitate the

sound of painting but the words ‘swish swish’ are used instead. This may be because the words ‘splash splash’ connote blots of paint, reflecting the impact of the paint on the surface, while the words ‘swish swish’ suggest a flow in the painting motion, reflecting the way the wrist flicks the brush up and down. It becomes apparent in Frame 7 (Figure 26) that the words ‘swish swish’ are most appropriate in this case as it is revealed that Naruto is painting spirals. The font used also corresponds with the linguistic representation. It can be described as ‘curly’ which evokes the idea that the sign being painted involves curves. It also suggests fluidity in the painting motion.



Frame 6

Figure 25: The notion of ‘us’ and ‘them’ is established through the foreground/background continuum (Kishimoto 2007: 10)

In Figure 25, Naruto is framed closer than before, even if only the bottom half of him can be seen. The point of view is from that of a third person who appears to be on the mountain face with Naruto. Even though the people below are given prominence as a result of the difference in colour saturation, the close proximity of Naruto encourages the reader to identify with him rather than the crowd. In addition, the notion of ‘light’ and ‘shadow’ evoked by the shading encourages the reader to direct his/her attention to the dialogues. However, as a result of the

reader's identification with Naruto, the shading creates the notion of 'us' and 'them' instead. The page ends with the frame 'bleeding' off the bottom of the page. On the next page, the consecutive frame 'bleeds' back into the page. This bleeding in and out creates a flow in the reading path. The two images appear to be a continuation of one another. It is also appropriate that the text has been structured so that the reader ends the page with a 'known' image (known in the sense that the image does not provide new information about Naruto), and begins the next page with a 'new' image (new in the sense that this is the first time the reader is presented a full view of Naruto). This echoes the information value which Kress and van Leeuwen (1996) attribute to the left and right side of a page except for the fact that the value is reversed in this case. On this page, the New is placed on the left and the Given is placed on the right. The right to left reading path in manga may play a role in the change in meaning.

In Frame 7 (Figure 26), Naruto swoops into view carrying a bucket of paint in one hand and a paint brush in the other. This is an extension of the image before. It is the first time that the reader is provided a full view of Naruto since the actual events of the narrative began so this first impression is important. Naruto's spiky hair emits an electrifying vibe, endowing him with a sassy look. This is a hairstyle usually associated with punks. The goggles on his forehead are an accessory and they evoke the idea of a little boy trying to be cool. The paint sloshed over him evokes a sense of grubbiness. His cheeky attitude is apparent through his facial expression and his taunting words. The narrator has been slowly building up a rebellious image of Naruto. The costumes and props here thus correspond well with this concept. Compared to the 'quiet' atmosphere in the last frame, this image is 'loud'. The jagged speech bubbles and the 'volume lines' suggest that Naruto is shouting. This brash atmosphere is reinforced by the fact that Naruto dominates the frame space.

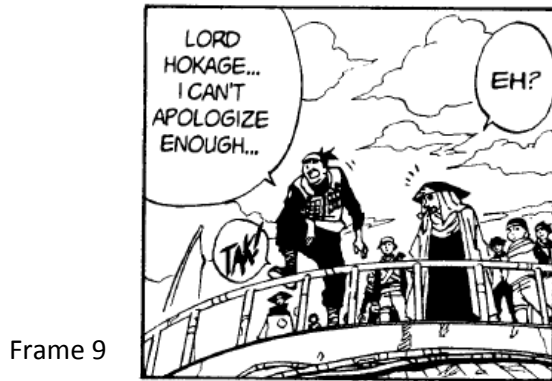


Frame 7

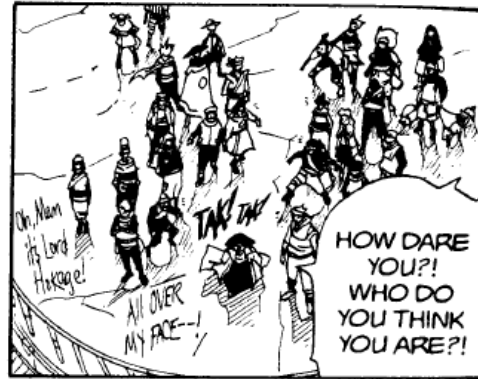
Figure 26: Naruto swoops into view (Kishimoto 2007: 11)

On the compositional metafunctional level, the absence of border lines on the margins of the page invites the reader into the narrative world. The co-deployment of camera angle, camera distance, frame and dialogue are significant in establishing a direct interpersonal relation. The frontal shot allows the reader to engage in the narrative. This is reinforced by the medium long shot which presents Naruto at a social distance. The framing of this shot along with the direct address of the dialogue invite the reader to interact directly with Naruto. It is as if the reader is part of the crowd.

In the next frame, Frame 8 (Figure 27), the narrative returns to the crowd below and the reader is afforded their reaction to Naruto's exclamation. This point of view is seemingly from that of Naruto's. In this image, Lord Hokage arrives to find that his face has also been defaced. He heads towards the edge of the stadium and as he walks, his footsteps "TAK! TAK!" can be heard even from above. This frame is, in fact, an extension of the image in Frame 7. The purpose is to reveal that Lord Hokage has joined the crowd.



Frame 9



Frame 8

Figure 27: Expanding an image through extension (Kishimoto 2007: 11)

The manner in which dialogue is employed in Frame 7 is particularly interesting. Generally, a speech frame only has a single ‘tail’ pointing towards a speaker. The function of the tail is to designate the speaker. In this case, there are three or four tails pointing randomly from the speech frame. This suggests that the dialogue comes from the general mass rather than a specific speaker. Interestingly, Lord Hokage’s exclamation “All over my face--!” and a ninja’s side remark “Oh, Man it’s Lord Hokage!” are not enclosed by a speech frame at all. This suggests that these are mutters that are not meant to be ‘heard’ by anyone in particular. The size and font of the mutters juxtaposed with that of the outburst from the crowd make the distinction between a ‘voice’ from an individual in the crowd and ‘voices’ from the crowd. This adds to the sense of ‘realism’ in the narrative.

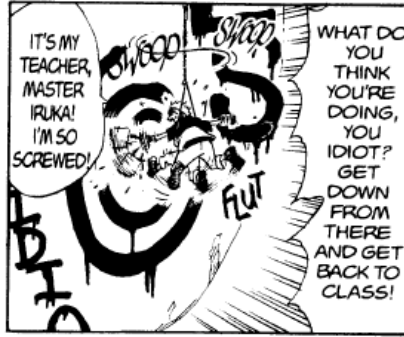
Baldry and Thibault point out that space is “*time-based*, that is, it is constructed around, and conditioned by, a *sequence of events* which involves the constant reorganisation of the participants’ occupancy of space in relation to each other” (2006: 6). This suggests that space has the ability to create the notion of time passing by depicting a represented element in one frame and then reconstructing their position in another frame. For example, Frame 8

illustrates Lord Hokage standing at a distance from the railing. In Frame 9, he appears to be nearer to the railing. This reconfiguration of space suggests that a moment in time has passed. The example demonstrates that space and time are closely bound in sequential visual narratives.

In Frame 9, Lord Hokage has reached his destination. However, before he had a chance to say anything, another ninja suddenly appears besides him out of nowhere. “TAK!”, with a foot on the railing, he address Lord Hokage and apologises for the situation. Although his speech is addressed to the lord his attention is fixed on Naruto. In this frame, the point of view has shifted from that of Naruto to a third person in the narrative. The angle has also changed – instead of an overhead shot, the angle is now positioned looking up at the characters. This suggests that the power relation has altered too. Before, the fact that Naruto dared to taunt the crowd suggests that he disregarded them. To the reader, the crowd was also merely a mass of people. No one stood out. Consequently, Naruto and the reader had power over the crowd. Now, Lord Hokage has appeared and the reader recognises him to be a man of importance. Although the ninja has not been introduced yet, by standing next to Lord Hokage, he too gains some respect from the reader. The camera angle reflects this change in attitude by being positioned at a low angle, looking up at the two characters.

With each image, the frame moves closer and closer to the subjects until finally in Frame 10 (Figure 28), the reader is placed at a close distance to Lord Hokage and the ninja. In Frame 10, Lord Hokage establishes that the ninja is “Iruka”. Iruka is given more prominence as a result of the space allocated to him. Despite this, the reader identifies more with Lord Hokage as the point of view is positioned from over his shoulder.

Frame 11



Frame 10



Figure 28: Expanding an image through projection (Kishimoto 2007: 11)

The last two frames illustrate an interesting case of extension through projection. Projection develops a sequence of images through dialogue (Matthiessen 2007). The dialogue can be internal as in thought or external as in speech. Figure 28 is an example of projection through speech. In Frame 10, the sound “SHF” along with Iruka’s body posture suggests that he is taking in a very deep breath. Frame 11 carries this action forward as his words burst through the speech frame. It is almost as if his body posture in Frame 10 serves as a catalyst for the pan to the next frame. In Frame 11, the intensity and volume of Iruka’s voice is expressed in the jagged speech frame. This corresponds well with Iruka’s big body movement. Despite the frames dividing the two images, the body posture and the projection allow a continuous flow in the reading path. Upon hearing this outburst, Naruto frets. This is visually represented by the ‘flut’ motion. As he flaps his arms and legs up and down, the rope holding him swings from side to side. This motion is represented by both the onomatopoeia ‘swoop swoop’ and by the curve line around the rope. Naruto’s muttering informs the reader that Iruka is in fact his teacher.

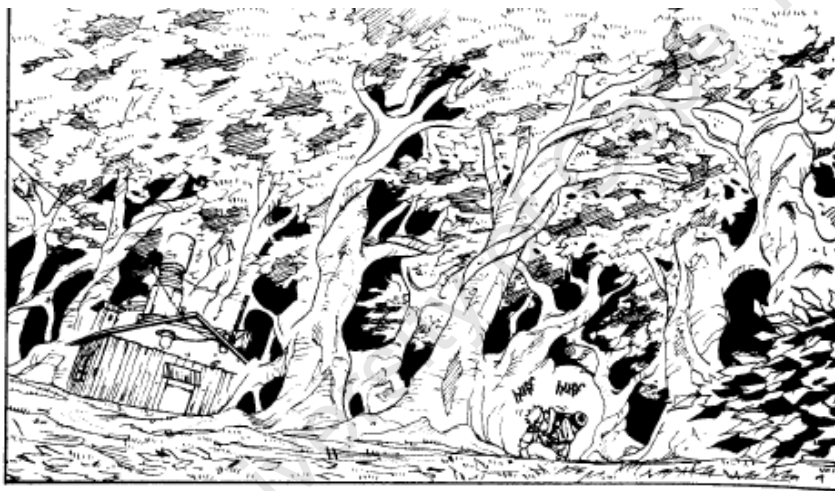
The positioning of the characters in the last three frames of this page is important to the flow of the narrative. The reader is introduced to Iruka for the first time in Frame 9. He is

presented on the left side of the image while Lord Hokage, a character with whom the reader is already familiar, is on the right side of the image. The Given/New value attributed to the left and right side of the page is reversed in this case due to the reading direction. The success of the projection is dependent on the fact that the image is read from right to left. This example demonstrates that information value is deeply rooted in the reading practices of a culture.

As mentioned before, Kress and van Leeuwen (1996) note that Eastern texts place greater emphasis on Centre/Margin compositions than polarised approaches to text composition. This seems to apply in manga too. In this sequence, the images are positioned at the centre while the speech frames are situated on the margins. This, however, does not mean that the images carry more weight than the dialogue. Since the reading path is left to right, readers are likely to follow this path when reading the frames. Instead, the Centre/Margin positioning of the images and written text evoke the sense that there is more balance in the 'functional load' (Kress 2003) carried out by the visual mode and the written mode. Functional load refers to the amount of information or meaning communicated by a mode (Kress 2003). Martinec (2003) uses the term 'communicative load' to refer to the same concept. In contrast, in Western comics, speech frames tend to be positioned at the top while the images are at the bottom. This positioning gives more weight to the written text as it is in the 'Ideal' position while the image is in the 'Real'. The written text is thus established as having supremacy over the image. This is reinforced by the fact that in Western comics, the written text tends to carry more weight than the image.

4.4 Complicating Action

The complicating action is the event which disturbs the balance established in the orientation and brings out a problem of some kind. In *Naruto*, this arises as a consequence of Naruto failing his graduation exam for the third time. This presents an obstacle to his goal of becoming the next Hokage in the village. Soon after he steals a forbidden scroll from the village headquarters and the ninjas pursue him. The following scene begins with the complicating action.



Frame 1

Figure 29: A wide shot re-establishing the location (Kishimoto 2007: 29)

The scene opens with a wide shot which re-establishes the location of the narrative (see Figure 29). The narrative has now shifted to the forest and from a distance, Naruto can be seen sitting and panting. On the experiential metafunctional level, the branches that twist around the trees create sense of eeriness. The place is desolate and mysterious. The low angle combined with the wide shot establishes the sense of a vast forest looming over Naruto as well as the reader. There is an overwhelming feeling of the power of nature. The angle is also

canted which creates the impression that the narrative world is off balance. This perspective is appropriate since the characters' world is about to be disturbed by the complicating action.

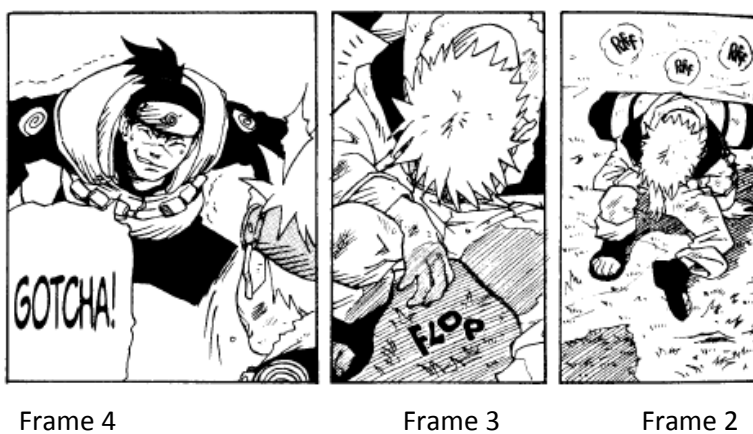


Figure 30: The use of elaboration creates tension in the narrative (Kishimoto 2007: 29)

The narrator elaborates on the scene by providing more detail on Naruto's circumstance. In Figure 30, Frame 2, Naruto is depicted at a social distance. He is hunched on the ground, panting heavily. The shadow in the left corner of the page suggests that somebody is approaching Naruto. The point of view is closely aligned with the anonymous individual. An elaboration of the image in Frame 3 brings the reader closer to Naruto. The shadow in the corner of the page grows larger in size. 'Flop', the sound of the individual's footstep alerts Naruto to his presence. Naruto's awareness of the approaching character is signified by the four dashes in the top left corner of the page. The point of view, the angle from which these two images are illustrated, along with the shadow which enlarges as the frame closes in on Naruto, creates the illusion of movement and of being present in the narrative. The reader seems to be approaching Naruto at the same time as the unknown character. A reverse shot in Frame 4 reveals the identity of this figure. It is Iruka, Naruto's teacher. The over-the-shoulder shot aligns this point of view with Naruto. The reader senses Naruto's surprise at being discovered through the lines around his face and the star-like shape above his head. The

reader also feels the tension building inside Iruka through his hunched body posture, forced smile, the drops of sweat on his face and the dotted lines around his body.



Frame 5

Figure 31: A moment of comic relief (Kishimoto 2007: 29)

The tension explodes in Frame 5 (Figure 31) and to accommodate for this explosion, the narrator takes a step back. With a medium long shot from a third person perspective, the reader witnesses the outburst from a social distance. The super-deformed depiction of a character is a manga convention and it conveys the sense of an extreme emotional state (Brenner 2007). Other symbols which accompany this image in expressing extreme anger is the pulsing vein, the streaks of lines on the forehead as well as the puff of smoke emitting from Iruka. In the events prior to this scene and even within this scene, the narrator has been building a level of tension. This dramatic representation of an outburst provides some level of comic relief within a sequence of serious events. The jagged frames which frame the speech of the characters display the force in their voices. This is reinforced by the ‘volume lines’. The reader can also experience the emphasis in their speech through the bold font used on certain words.



Figure 32: Close-up shots depicting reaction after outburst (Kishimoto 2007: 29)

Frames 6 and 7 (Figure 32) are reaction shots following the outburst. It is appropriate that Iruka's restored state is shown straight after his deformed state. This creates an impression of a 'before and after' experience. In addition, by having the images follow one another, there is less of a disruption to the flow of the narrative. In Frame 6, once again, the emotions are symbolically represented through shapes and lines. The dotted frame encompassing an exclamation mark is neither a speech nor a thought frame. It appears to be a comment on Iruka's internal emotion. Baldry and Thibault use the term 'cluster' to refer to "groupings of resources that form recognisable textual subunits that carry out specific functions within a specific text" (2006: 11). Speech and thought frames are examples of clusters. The cluster here signifies Iruka's emotion, that he is still irked by Naruto's actions. This is reinforced by the puff of cloud which signifies 'letting off steam'. From this, it transpires that shapes and lines can convey both symbolic and interpersonal meanings. It is appropriate that in the positioning of these frames, Frame 6 and 7, Iruka is positioned above Naruto. Iruka is not only taller than Naruto but he is also the teacher. It is therefore fitting to have him positioned higher than the student. This suggests that layout is also capable of signifying social relations.



Frame 1



Frame 4

Frame 3, Frame 2



Frame 6

Frame 5



Frame 7

Figure 33: Frames establish tempo (Kishimoto 2007: 29)

In sequential visual narratives, different pages afford different reading duration as a result of the size and the number of frames used to depict the narrative. The different reading duration establishes a tempo in the narrative. This is a narrative device which allows the narrator to work towards a climax. Take the frames on this page for example (see Figure 33). The first frame is large – its size invites the reader to linger on the image. This is reinforced by the timeless presence established through the ‘bleed’ effect. The feeling changes in the second line of frames. On the second line, the frames are smaller and more or less uniform. In fact,

Frames 2 and 3 are the same size. This creates a staccato in the tempo. The staccato speeds up the pace and at the same time builds up tension in the narrative. Frame 4 allows for a short pause because the frame is slightly larger and the reader has to stop to read the dialogue. The tension, however, is still present and is reinforced by the represented elements inside the frame. In Frame 5, as the tension boils over, the size of the frame expands to match the action inside the frame. As the reader reads the dialogues between the two characters, a long break is established. This is the moment of climax so the long pause is appropriate. In the last two frames, the pace quickens again to prepare the reader for the actions on the next page. From this it is evident that frames are important in establishing the tempo of the narrative and the reading pace. It also becomes clear that the pace of the narrative and the size of the frames correlate with the length of the dialogue.

After a brief comic moment, the narrative resumes to the event which culminates in the complicating action. Once again, the turning of the page acts as a transitional device. This time it assists in the transition of the mood of the narrative, moving from the comic atmosphere established in Frame 5 back to the serious tone of the complicating action.



Frame 8

Figure 34: Ominous mood established through framing (Kishimoto 2007: 30)

The experiential metafunction in Frame 8 (Figure 34) is established largely through the angle and body posture of the represented elements. The narrator restores the intensity established at the beginning of the complicating action by depicting the narrative world from a rather provocative point of view. The low angle sets an ominous mood as the reader is required to engage in the narrative world with the characters, especially Iruka, towering over. The sense of an impending disaster is reinforced by the black ‘cloud’ hovering at the top of the page. Iruka’s role as an authority is established by his body posture. With his hands on his hips, he is posed to reprimand. His authority is reinforced by his size in proportion to Naruto. A teacher-student relationship becomes apparent through the angle and the body posture employed.

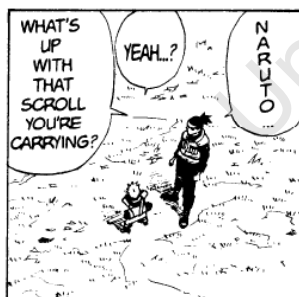
The village ninjas assumed that Naruto had stolen the forbidden scroll as a prank. Naruto reveals that this was not the case in Figure 35, Frame 9. He took the scroll in order to learn the skills to graduate. His excitement is conveyed in the jagged edges of his speech frames and ‘volume lines’. The close-up view allows the reader to feel the intensity in his enthusiasm. Upon hearing this, Iruka is surprised. The black background in Frame 10 suggests that this shock is felt in his internal world. The star-like sign indicates a moment of revelation.



Figure 35: A moment of revelation (Kishimoto 2007: 30)

The move from Frame 10 to 11 is another example of extension through projection. In this case though, the projection is through an internal dialogue. The black background in Frame 10 establishes that the reader is accessing Iruka's internal world. In Frame 11, the words unmistakably belong to Iruka but they appear without speech frames. This suggests that the words and point of view belong to Iruka. If Frame 10 had been omitted then the flow in the narrative would have been less smooth as the reader would have to jump from one character's dialogue to another character's thoughts. Frame 10 thus functions as a linking frame which creates a channel for the reader to enter Iruka's internal world and access his thoughts and point of view.

To demonstrate that the narrative has exit Iruka's internal world, the next image, Frame 12 (Figure 36), illustrates the narrative from a long angle and an omniscient point of view. These two resources clearly establish that the reader has exited Iruka's internal world and re-entered the narrative as someone outside it.



Frame 12

Figure 36: An omniscient point of view (Kishimoto 2007: 30)

In Figure 37, Frame 13 and 14, Naruto reveals the cause of the complicating action. According to Naruto, Master Mizuki, another teacher in the academy, told him that in order to graduate he needs to demonstrate to Iruka that he can use the techniques in the scroll. He

evidently does not know that the scroll is forbidden and that he is in deep trouble for taking it. Instead, he is excited that Iruka has discovered him and thrilled that he will have another chance to prove himself. In Frame 13, this excitement is suggested in the white jagged ‘aura’ that seems to emit from Naruto. In contrast, in Frame 14, the excitement is represented by the volume lines. In language, synonyms replace one word or phrase with another. From this example, it would appear that there are ‘synonyms’ in images too. Despite this, there are reasons why individuals choose to use certain words over others. As Kress points, the sign is “always both a representation of what it was that the sign-maker wished to represent, and it is an indication of her or his interest in the phenomenon represented *at that moment*” (2003: 144). Thus, the question is what is the motivation behind using a black background in the one image and a white background in the other?

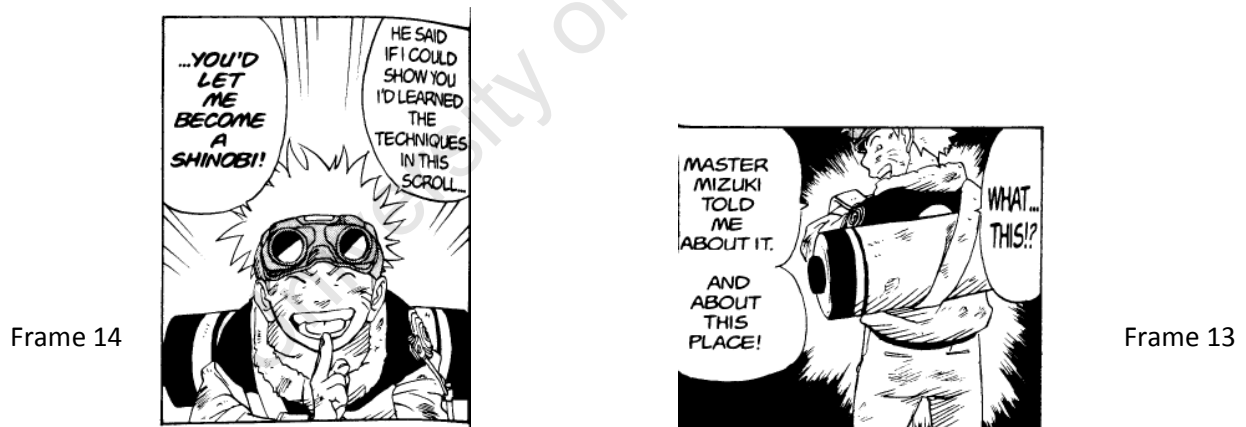
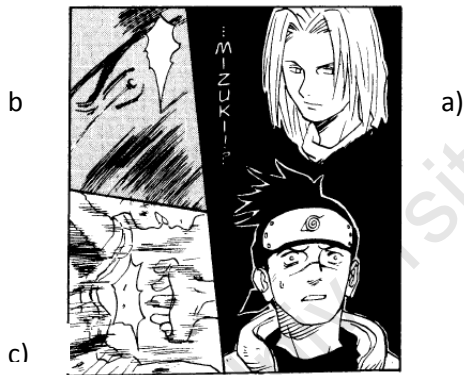


Figure 37: The cause of the complicating action revealed (Kishimoto 2007: 30-31)

In Frame 13, the white jagged frame evokes the sense of excitement but this excitement is enclosed by a black background. This seems to suggest that the emotion is controlled. In Frame 14, the white background is pervasive. This implies that Naruto is elated. This notion is reinforced by the bold font used in the dialogue. Linguistically, these two frames can be

expressed as: Excitement builds inside Naruto as he begins his recount. By the time he finished his recount, he is ecstatic. It is also appropriate that the background should change from dark to light as Naruto flips from back to front. The colour change accentuates the position change.

Martinec and Salway propose that “some kinds of images may be better at creating direct emotional impact, and text may be more suited to carrying out logical analysis” (2005: 338). From this sequence of images, it surfaces that the written mode is mostly used to summarise and provide the logical meaning in the narrative while the visual mode is used mostly to present the experiential and the emotional meaning.



Frame 15

Figure 38: A split frame (Kishimoto 2007: 31)

In Figure 38, Frame 15, upon hearing Mizuki’s name, Iruka freezes and a drop of sweat rolls down his face. His facial expression is tense. This indicates that Iruka senses some problem with Naruto’s recount. In Frame 15a (Figure 38), the black background once again indicates that the narrative has shifted to Iruka’s internal world. But he does not stay there for long. He is brought out of his thoughts abruptly by a movement in the external narrative world. This unexpected snap back into the external world is portrayed by the star-like sign. Frame 15c

depicts the movement of someone throwing something in the air. These actions are represented in one frame which suggests that they happen within seconds of each other.

Figure 39 (Frames 16, 17 and 18) illustrates a series of action. In Frame 16, the reader is confronted with daggers flying through the air. This is a subjective point of view but at this instance, it is not clear whose perspective it is. In Frame 17, the viewpoint shifts to that of a third person point of view. This frame reveals Iruka as the target of the daggers. He is pushed back by the blade and slides to a stop in Frame 18.



Figure 39: Expanding a sequence of images through extension (Kishimoto 2007: 31)

An approach to representing motion in static images is through the use of kinetic lines often referred to as 'speed lines' or 'motion lines'. These lines make it possible to create the illusion of movement within a single frame. In employing motion lines to depict movement, manga uses an approach that is different to Western comics. As depicted in these images, the streaked lines create the illusion that the reader is moving with the character. McCloud terms

this “subjective motion” as it is meant to provide the reader a subjective experience (1994: 114).

The images in this sequence are developed through extension. That is, each of these images expands one another by adding some new but related information. In Frame 16, a rain of daggers flies towards the reader. The point of view is subjective so it leaves one to question whose perspective has the reader adopted? Frame 17 provides the answer to this by shifting to a different viewpoint. The new point of view allows the reader to experience the action from a third person perspective and this consequently provides more information with regard to the context of the situation in the narrative. That is, Iruka is under attack and the point of view in Frame 16 belonged to him. Frame 18 signifies the end of the attack as he comes to a screeching stop. Thus, one image adds to the meaning of the other, allowing the reader to put the pieces together into a coherent whole.

The extension here shows aspects of an action happening over a quick moment in time. The images only depict parts of the action yet it is possible for the reader to piece together the parts to form a coherent narrative. This phenomenon can be explained through McCloud’s (1994) concept of ‘closure’. According to McCloud, closure is “the phenomenon of observing the parts but perceiving the whole” (1994: 62). What this means is, for example, if we see a pair of feet, then logically we will assume a body is attached to the feet despite not actually seeing the body. Closure allows the reader to follow a narrative moving from frame to frame without the artist drawing every moment of the action. It is “the agent of change, time and motion” in the narration (McCloud 1994: 65). For closure to work and the narrative progression to be successful, the reader’s imagination and participation is extremely important. The degree to which the reader has to work to piece together frames of images to

form a coherent narrative differs though. For example, in this case, a high reader involvement is necessary as only parts of a sequence of action are illustrated. The reader is required to construct a coherent flow of action by seeing on parts of it. Figure 40 presents a complete layout of the complicating action.



Figure 40: Complicating action (Kishimoto 2007: 30-31)

In the events that follow, Mizuki emerges as the villain in the story. He wanted the scroll for himself and set Naruto up to steal it for him. It would have been easy for the village of ninjas to blame Naruto for the deed since he is generally disliked. He not only frequently causes trouble through his mischief but is also loathed because he is often taken for the nine-tailed demon fox itself. As mentioned in the abstract, the demon fox caused the villagers great

suffering. It took much effort to capture the demon and although the ninjas succeeded in the end, they lost their revered Hokage to the deed. Now, it emerges in the narrative that as a baby, Naruto, had the demon fox sealed inside him in order to prevent it from completely destroying the village. Tragically, instead of being regarded as a hero by his fellow villagers, he is considered to be the demon fox itself, leading him to be ostracised from society. The villager's prejudice against Naruto creates the perfect opportunity for Mizuki to set him up as the villain.

4.5 Evaluation



Figure 41: Evaluation (Kishimoto 2007: 48)

The evaluation provides an assessment of the events and functions to disclose the purpose of the narrative. In Labov's words, evaluation is "the means used by the narrator to indicate the point of the narrative, its *raison d'être*: why it is told, and what the narrator is getting at (Labov 1972: 366). The evaluation here (see Figure 41) reveals that this is a narrative about an individual's struggle for recognition. This is a story about Naruto's struggles to be recognised as Naruto, a ninja of indispensable value, rather than the demon fox threatening the well being of the village.

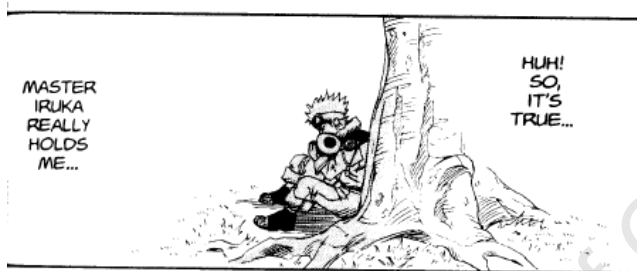
Prior to this sequence of images (Figure 41), Iruka had told Naruto to hide and to safeguard the scroll. In this sequence of events, the reader finds Naruto hiding behind a tree. From his hide out position, he overhears a conversation between Iruka and Mizuki (see Figure 42).



Figure 42: Naruto overhears a conversation between Iruka and Mizuki (Kishimoto 2007: 48)

In Figure 42, Frame 1, the image is framed from a third person point of view but the reader identifies more with Iruka because he is placed at a closer distance to the reader than Mizuki. Frame 2 and 3 reveals Naruto's shock and anger at hearing that Iruka had the same sentiments about him as the rest of the villagers – that he too despised Naruto. The emotional

impact of Naruo’s shock and anger is vivid in Naruto’s facial expressions. In Frame 2, the notion of shock is suggested by Naruto’s open month and wide eyes. He is clearly baffled by Mizuki’s words. This is reinforced by the star-like shape on the upper corner of the image which signifies surprise and a moment of revelation. The black background in Frame 3 signifies the transition into Naruto’s internal world. In Frame 3, his face is scrunched up in anger. This is reinforced by the letters “GRRR” which suggests that he is growling. The bold font of the letters signifies the intensity of the emotion.



Frame 4

Figure 43: The white background signifies Naruto’s isolation (Kishimoto 2007: 48)

The anger turns to a sense of loss and loneliness in Frame 4 (Figure 43). The whiteness or ‘nothingness’ in the background of this frame reflects Naruto’s isolation and loneliness. This is reinforced by shading Naruto in grey. The colour signifies dismay. In Frame 5 (Figure 44), he literally and figuratively plunges into complete darkness as he realises that he has no one who cares for him. The entirely black background is especially powerful at this moment in the narration. It firmly establishes Naruto’s despair and hopelessness. However, just as Naruto was about to fall into a state of gloom, Iruka’s voice snaps him out of the dark internal world. Iruka’s speech bubble overlaps the frames and this suggests that the dialogue carries across both frames. The overlapping speech bubble thus functions as a bridge connecting one frame with another.

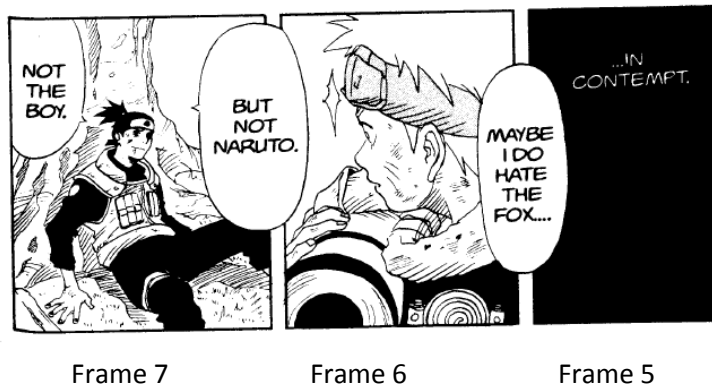
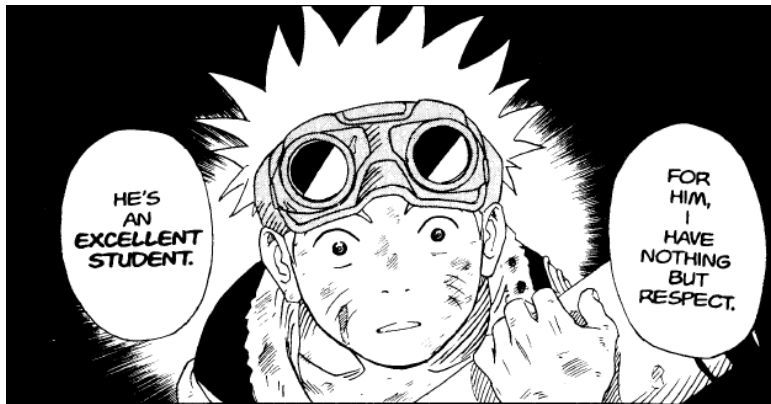


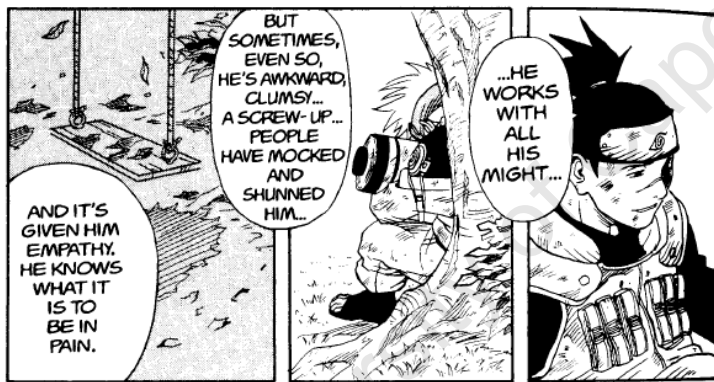
Figure 44: Speech that overlaps frames (Kishimoto 2007: 48)

In the sequences of images that follow (Figure 45), the narrative is propelled forward through projection. At the same time that Iruka's words pulled Naruto back into the narrative world, the words become the focal point of the narrative. The reader, like Naruto, is interested to find out what Iruka means when he says that he hates the fox but not Naruto. At this point, the words are important as they provide the logical meaning to the narrative. The evaluation of the narrative is explained through Iruka's dialogue. Since his speech is spaced out over a number of frames, it propels the sequence of images forward. Iruka's words, mostly appearing in the form of voiceovers, also functions to bridge the seemingly disconnected images, in particular, Frame 10 and 11. The move from Frame 10 to 11 can be described as transition since Frame 11 is set in another place and at another time. The grey shading suggests that this image is a 'memory'. Without the voiceover connecting the images, the transition would have been abrupt as the empty swing is out of context in this sequence of events. Nevertheless, the frame is very important in providing the experiential and the emotional meaning. The empty swing and the falling leaves strongly evoke the sense of sorrow and loneliness. From this, the reader can vividly sense the dejection of being an outcast. The image functions as a visual metonym as it acts as a substitute for the various emotions associated with being an outcast.

Iruka's evaluation of Naruto here is the turning point of the story as he presents a view of Naruto which is completely different to that of the villagers.



Frame 8



Frame 11

Frame 10

Frame 9

Figure 45: Developing the sequence of images through Projection (Kishimoto 2007: 49)

McCloud notes that manga tend to place more emphasis on “being there over getting there” (1994: 81). In other words, manga tends to pace out the narrative and wander more, placing more emphasis on the experience of the narrative rather than getting to the point of the story.

In fact, McCloud notes that this is a trait of Eastern art and literature as a whole.

Traditional Western art and literature don't wander much. On the whole, we're a pretty goal-oriented culture. But, in the East, there's a rich tradition of cyclical and

labyrinthine works of art. Japanese comics may be heirs to this tradition, in the way they so often emphasize being there over getting there” (McCloud 1994: 81).

Martinec (2003) notes the same phenomenon in his analysis of Japanese recipes. In comparing Japanese recipes with English recipes, Martinec found that the images in English recipes tend to show the finished product while in Japanese recipes the images portray the different stages of the cooking process. A similar case can be noted in this sequence of images (Figure 46).

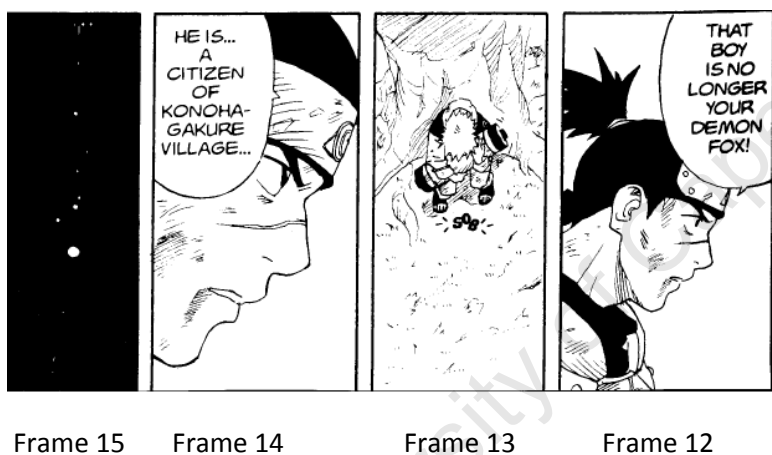


Figure 46: Pacing the narrative through wordless panels (Kishimoto 2007: 49)

The evaluation could have been dealt with in a more condensed manner, but the author paces the narrative in order to draw out the tension in the event. In particular, the wordless panels could have been omitted but they are deliberately positioned at particular points in the sequence in order to pace the narration and at the same time to draw out the tension in the narrative. The effect of the wordless panel in Frame 15 (Figure 46) is notably powerful.

Frame 15 is prominent because the black background is in sharp contrast to the white background of the other frames. The blackness also establishes a sense of complete silence

which in turn creates an atmosphere of suspense. Since the reader can see every teardrop falling, it suggests that the action is in slow motion and this adds to the suspense. Besides establishing suspense, the teardrops also provide a poetic sense to the narrative. Considerable effort is thus made to provide the reader with the experience of being present in the narrative. McCloud suggests that it is this “amplified...sense of the reader participation in manga, a feeling of being part of the story rather than simply observing the story from far” which makes manga so captivating and popular (2006: 217). It can thus be said that there is a high level of engagement in manga. The engagement is not only realised by pacing out the story in order to elicit the mood but also through what Martinec (2003) describes as the ‘system of engagement’. By ‘system of engagement’, Martinec refers to “the degrees of interpersonal closeness or distance realised by a combination of body distance and angle between interactants” (2003: 46). This is evident in Figure 47 where the sequence of images culminates in a burst of emotion.



Figure 47: Climax (Kishimoto 2007: 50)

The emotional intensity of this image owes itself to the strong degree of engagement realised through the size of the frame (Figure 47 takes up half of the space on this page), the close-up angle which establishes a strong interpersonal closeness and importantly, the powerful facial expression. This is the point of climax as Naruto realises that there is someone who finally accepts him for who he is, as Naruto, an individual, rather than a demon fox. Naruto's facial expression and gesture evoke a strong emotional experience. With a clinched hand and tears pouring down his face, the facial expression and gesture appear to signify extreme gratitude at being acknowledged. It seems that all the tension and loneliness which had been hidden inside Naruto his entire life is released at this point. The effect of this climaxing point is reinforced by the suspense created in the last frame, the 'pause' established by the page turning and the composition of the image.

4.6 Resolution

The resolution closes the sequence of complicating action by solving the problem one way or another. For a while, it seemed like Iruka would die in the hands of Mizuki. Naruto's sudden appearance at the scene of conflict, however, changes the direction the story was taking (see Figure 48).

It is interesting to note how various resources are employed to represent the sense of time and action in Figure 48. A vivid sense of time and duration in time is constructed by the various frame sizes and the space they occupy. In Figure 48, Frame 1 overlaps Frame 2. This creates the idea of 'at the same time'. The duration of the actions is implied by the size of the frames. The smaller frame suggests a quick moment in time, while the 'bleed' effect of the larger frame creates an extended sense of time. The bleed frame encourages the reader to linger on

the action and fully experience the movement. Action in this fight scene is represented through motion lines, the blurring of the characters and body posture. In Frame 2, Mizuki is positioned for a forward movement. The illusion of this movement is propelled forward through the combination of the motion lines and the blurring of Mizuki's figure. Likewise, in Frame 7, Mizuki's body is positioned for a fall and this action is aided by the motion lines. The actions depicted here are extremely important in establishing the experiential metafunction. *Naruto* is a 'shonen' manga, a category of manga intended for boys, therefore fight scenes are expected. The motions lines and sound effects are foregrounded in this sequence because they are resources which best convey action.



Frame 1

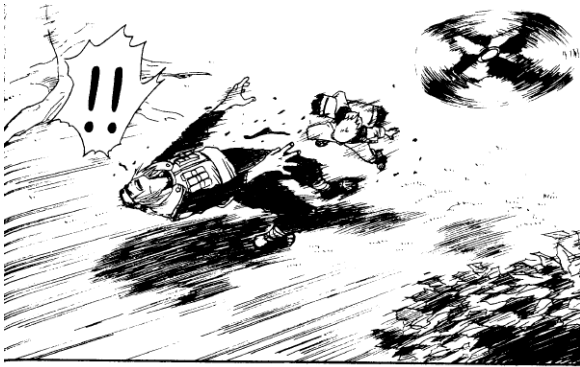
Frame 2

Frame 6

Frame 5

Frame 4

Frame 3



Frame 7

Figure 48: Actions prior to Resolution (Kishimoto 2007: 51-52)

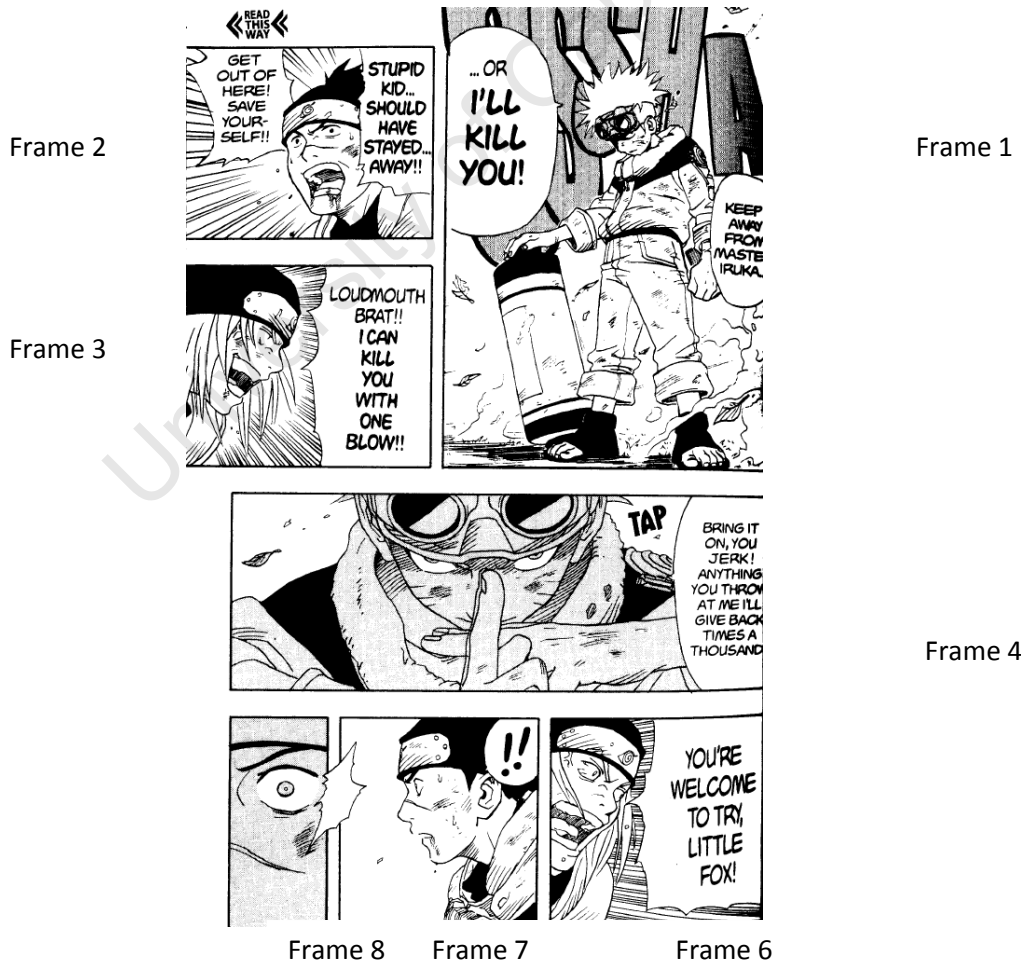


Figure 49: Resolution (Kishimoto 2007: 53)

Naruto's sudden appearance at the scene of the conflict signals the end of the complicating action and the beginning of the resolution. The resolution (Figure 49) presents a transformed Naruto – a Naruto that is confident, serious and threatening. All these adjectives describing the transformed protagonist are expressed through a combination of semiotic resources. For instance, in Frame 1, the sense of dominance and threat are established through Naruto's body posture and the low camera angle employed. The feeling of intimidation is reinforced by the words **"I'LL KILL YOU!"** in bold. The reader experiences this threat first-hand as Naruto glares directly at the reader in Frame 4. The effect of this fierce look is reinforced by the grey shadow cast over Naruto. Brenner (2007) points out that this is a manga technique signifying extreme emotion. In this case, the grey signifies a grave, sombre feeling. Naruto's hand gesture at this point indicates that he is about to perform a ninja technique. Judging by Iruka's reaction in the last two frames, the result of this technique is astounding. In Frame 8, Iruka's face literally changes colour from shock – the grey shading signifies great astonishment. The Naruto presented here is one that the reader has not yet encountered before. It is appropriate that this transformation signalling character growth should come at a stage where there is a turn in the events of the narrative.

The extreme close up of Iruka's reaction to Naruto's ninja technique cues the reader for some unbelievable action in the next frame and indeed, the action is unbelievable. As the reader turns over the page, what seems like a thousand replicas of Naruto explodes before the reader (see Figure 50). The extreme intensity of this image derives from the use of the entire double page to illustrate the one image. Naruto is literally everywhere. The bold typography matches the audacity of the image. Ironically, the 'Art of the Doppelganger' is Naruto's worst skill. He failed three times at the ninja academy because he could never produce a solid replica of himself. Yet, in this instance, he is able to multiply himself by an overwhelming amount.

This signifies his growth as a character. From henceforth, Naruto will no longer deliberately cause trouble. He will protect those he loves and be a hero of the village.



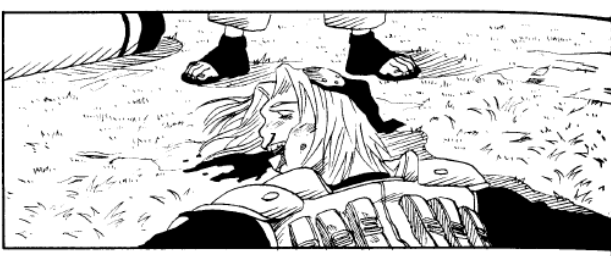
Frame 9

Figure 50: 'The art of the doppelganger' (Kishimoto 2007: 54-55)

The function of a resolution is to solve the problem that caused the complicating action and answer the question 'what finally happened?' In Frame 9, the cause of the complicating action (Mizuki) is minimised to a tiny entity at the centre of the page. The problem literally shrinks as the protagonist releases his full potential. So what finally happened?



Mizuki gets beaten to a pulp...



...and Naruto graduates.



4.7 Final comments

This chapter has taken a social semiotic approach to analysing *Naruto*. The aim is to demonstrate how various semiotic resources are used in manga to narrate a story. It emerges from this analysis that each semiotic resource used in the narrative has distinct story-telling functions. The next chapter presents an overview of the narrative functions of the various resources employed, as well as draws out the implications of the analysis. The second research question “what are the possible implications of using a metalanguage in teaching other visual narratives?” is thus addressed.

University of Cape Town

Chapter Five: The Implications of the Study

5.1 Overview of chapter

This chapter draws out the implications of the study. The first half of the chapter identifies the semiotic resources explored in the previous chapter and presents an overview of their narrative functions. It includes a discussion on modes and logics and highlights the influence of social and cultural factors on conventions of manga. The second half of the chapter addresses the research question: “what are the possible implications of using a metalanguage of manga in interrogating other visual narratives?” The chapter proposes a scenario where a metalanguage of manga can be used to interrogate storyboarding. The chapter closes with an overview of the contributions of the study.

5.2 Semiotic resources and their affordances

In chapter four, manga was analysed from a multimodal social semiotic perspective in order to disclose how various semiotics resources can be used to recount a visual narrative. Harvey argues that “[i]n the best, the pictures do not merely depict characters and events in a story: the pictures also add meaning-significance to a story” (1996: 3). Indeed, it emerges from this analysis that images have important narrative functions. It becomes clear that different semiotic resources have different meaning-making potentials and each contribute to the narrative in different ways.

The frame, for instance, emerges as an important narrative device in expressing time and cause and effect. These are fundamental characteristics of stories. As Baldry and Thibault

point out, narratives “do not merely signal a temporal succession of events. More importantly, they show how some aspect of a situation or a participant in a narrative *changes* as a result of the transition from an earlier moment to some later moment” (2006: 13). Nevertheless, still images are only capable of illustrating one moment in time which means they are limited in terms of expressing causality and temporality. In language, Kress notes that “sequence of events as represented in sequence of clauses is often open to a causal interpretation” (2003: 57). Baldry and Thibault too comment that “[t]he very notion of sequence implies a time-based, chronological ordering of events in a narrative and/or cause-effect structuring” (2006: 44). Thus a strategy of evoking causality and temporality in still images is to employ frames in sequences. Sequential frames make it possible to express the passage of time and cause and effect by dividing a narrative event into specific moments. By connecting the various moments depicted in the frames and through the concept of ‘closure’, the reader can infer the sense of a narrative progressing.

Another meaning-making potential of the frame is its ability to evoke a sense of duration in time through a manipulation of the size or the borders of the frame. It transpires from this analysis that a frame protruding off the edges is often able to convey a sense of timelessness, while a small frame suggests an instant in time. The length of time conveyed establishes a certain pacing or tempo within the narrative. Besides these functions, frames can also serve as a metacomment on the world of the represented. As in the case of the abstract (Figure 16), by enclosing the narrative world in a frame and distinctively separating it from the ‘outside’ world, it signifies that the world enclosed by the frame differs from the world outside and therefore should be regarded differently. The frame can also communicate interpersonal meanings. In this analysis, this emerges mostly in the case of speech frames. A jagged speech frame can suggest great excitement while a smooth one can signify a relaxed speech. It is

important to mention that the interpretation of these interpersonal meanings is dependent on the context of situation in the narrative.

Colour is another semiotic resource with vast meaning-making potential. Except for the front cover, this manga narrative is realised in black and white. This is the case with most manga narratives. Monochromatic images are usually perceived as a cheap form of art and indeed, this is a factor in manga's relatively low cost production. However, it emerges that even under such a constraint, the two tones are capable of conveying meanings on three metafunctional levels.

On the representational metafunctional level, colour emerges as being able to signify specific spaces. On a number of occasions in the narrative, a black background is used to show entrance into a character's internal world while a white background adjacent to the image signifies a return to the external world. The colours black and white therefore become important signifiers of the internal world and external world, of thought and 'reality'. Colour is also capable of enhancing the experiential meaning. Different shades of grey, for instance, create the illusion of 'shadow'. This provides depth to the image and adds to the sense of realism.

On the interactive metafunctional level, colour emerges as able to express interpersonal meanings. For example, a grey shadow cast over a character's face or figure is can signify gloom or grimness. Used with other semiotic resources, the juxtaposition of black and white can also convey excitement (see Figure 51). In Figure 51, Naruto's excitement is evoked by the 'energy' that seems to emit from him. This 'energy' is represented by the colour white and the jagged lines around the edges.

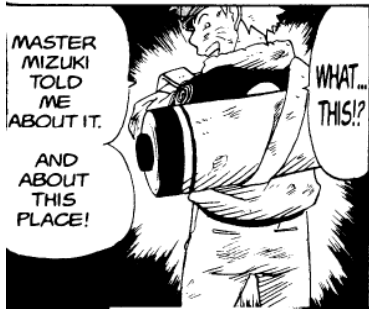


Figure 51: Illustrating excitement through colour (Kishimoto 2007: 30)

On the compositional metafunctional level, colour emerges as being able to attract viewer's attention through contrasts in the tonal value. The establishing shot in the orientation (Figure 4.4) is a prime example of where shades of grey on the margins guide the reader's attention to the white space at the centre of the page. Once again, it is important to note that the implied meanings of these uses of colour are derived from the context of situation in the narrative. In most cases, other semiotic resources are often necessary to assist in the interpretation of these meanings. This points to the importance of reading a text as an integrated whole.

Layout is another important device in visual narratives. The analysis of the Japanese and English edition of *Naruto* demonstrates that layout greatly influences the reading experience of a narrative. The composition in the Japanese version evokes a sense of mystery while the English edition conveys a sense of authority. On the whole, it emerges that centre/margin composition creates a sense of balance and harmony while top/bottom composition evokes the sense of hierarchy.

The layout of the text also affects the flow of the narrative. Scene transitions and climax points are more effective when placed in certain positions on a page. In this analysis, page turning emerges as an important transitional device, effective in helping scene changes and

achieving moments of climax. It also becomes apparent that the information value of Given/New attributed to the left and right sides of the page are reversed in manga. Jewitt and Oyama (2001) note the same phenomenon in their comparative analysis of British and Japanese advertisements. It becomes evident that the values attributed to positions of a page are culturally and socially situated. It is important to recognise that there are values attached to various positions of a page in all texts (Kress and van Leeuwen 1996). The key is to read these values in the context of situation and context of culture from which the text emerges. Having said this, Kress and van Leeuwen note that “signifying systems will have relations of homology with other cultural systems, whether religious, philosophical or practical” (1996: 199). This suggests that meanings attributed to positions of a page are likely to overlap in many cases. For example, in this analysis, the information value attributed to the centre and margin and top and bottom of the page do not appear to differ from that of Western values. The key difference is in the information value assigned to the left and right side of the page. This study attributes this disparity to the difference in the reading practices of the East and West.

The human form, this includes facial expressions, body posture and gesture, is a powerful resource in signifying interpersonal meanings. Facial expressions are especially effective when it comes to portraying emotions. For example, in Figure 47, the climax moment in the evaluation, Naruto’s expression and gesture conveys a burst of emotions – a sense of gratitude, a sense of being deeply moved. The meaning of the image is polysemous and it is also relatively open to the reader’s interpretation. Facial expressions, body posture and gestures tend to be exaggerated in manga but this amplifies the sense of reader participation as he/she comes to identify with the characters, their actions and emotions. According to McCloud, “[h]umans love humans! They can’t get enough of themselves. They crave the

company of humans, they value the opinion of humans and they love hearing stories about humans” (2006: 60). It is thus not surprising the human form is most effective in drawing emotional responses from readers.

Point of view, social distance and angle positions are semiotic resources which are always realised together and they are important in establishing interpersonal meanings. Point of view and social distance control the amount of information communicated to the reader. A subjective point of view provides the reader with a first-person experience but it limits the amount of information communicated. Likewise a close-up provides a detailed view of the represented object but conveys very little about the context of situation. On the other hand, a third person or an omniscient point of view combined with a wide shot provides more insight into the context of situation. Angle positions, as noted by Kress and van Leeuwen (1996), are capable of disclosing attitudes and levels of engagement. It also becomes apparent from the analysis that the positioning of the angle can comment on how the represented world should be viewed. For instance, a canted angle could imply a distorted world.

Lines and shapes are other semiotic resources employed which help to construct meaning in visual narratives. Lines can function as vectors which direct the reader’s attention to participants. They can evoke the loudness of the dialogue through ‘volume lines’. In addition, they can represent motion. Both lines and shapes can express meaning on an interpersonal level. Read in conjunction with other semiotic resources, they can signify emotions such as excitement, anger or tension. For example, in Figure 52, the dotted lines which frame the exclamation mark convey tension and alarm. The puff of smoke suggests that Iruka is ‘letting off steam’.



Figure 52: Lines and shapes as semiotic resources (Kishimoto 2007: 29)

Typography is an important device in paper-based narratives. It functions as “a transmitter of the written word” (van Leeuwen 2004: 14) and is capable of communicating meaning on a sensory level. From this analysis, it emerges that typography is an important resource in representing voice and sound effects. As a result of the font, lettering and other graphic elements which form part of the typographical design, it is possible to convey tonal inflections, volume and even the timbre of sound. For this reason, McCloud notes that words in comics provide “readers a rare chance to listen with their eyes” (2006: 146). The graphic design made possible through typography renders the concept of ‘homospaciality’, the co-deployment of two semiotic modes in one unit (Unsworth 2006), a common occurrence in comic art.

The metalanguage is an important resource which has made it possible to identify and describe the various elements of the narrative and their semiotic potential in this study. Using the metalanguage, this analysis has demonstrated that each of the semiotic resources used in the narrative communicate distinct meanings and contribute towards the reading experience on different levels. The meaning of the narrative, however, is conveyed from an integration of all the resources not on the basis of individual semiotic resources. Baldry and Thibault

(2006) use the term ‘resource integration principle’ to refer to the necessity of viewing texts as multimodal.

In practice, texts of all kinds are *always* multimodal, making use of, and combining the resources of diverse semiotic systems in ways that show both *generic* (i.e. standardised) and *text-specific* (i.e. individual, even innovative) aspects (Baldry and Thibault 2006: 19).

Moreover, meaning always emerges as a result of the integration of semiotic resources.

Multimodal texts integrate selections from different semiotic resources to their principles of organisation...These resources are not simply juxtaposed as separate modes of meaning-making but are combined and integrated to form a complex whole which cannot be reduced to, or explained in terms of the mere sum of its separate parts (Baldry and Thibault 2006: 18).

The resource integration principle is clear in this study. It is not possible to tell the story of Naruto, for instance, through the use of sequential frames alone. Other modes such as writing and images are necessary. To give another example, sequential frames are noted to convey temporality. However, the notion of time cannot be expressed without a change of state in the represented elements captured in the frames. Time is therefore not conveyed through the frames alone but through a combination of resources. The metalanguage is key to identifying the semiotic resources and providing the language to discuss how they function to make meaning.

5.3 Mixing logics

Kress (2003) argues that the visual mode and the written mode are governed by different logics. However, it emerges from this analysis that this distinction is not too clear cut.

According to Kress,

The organisation of writing – still leaning on the logics of speech – is governed by the logic of time, and by the logic sequence of its elements in time, in temporally

governed arrangements. The organisation of the image, by contrast, is governed by the logic of space, and the logic of simultaneity of its visual/depicted elements in spatially organised arrangements (2003: 2).

In manga, writing is treated as a visual entity and therefore governed by the logics of space. That is, writing occupies space and its position in this space determines its value and the sequence in which it will be read. The visual entity, on the other hand, is treated as a ‘written entity’ on some levels. This is as a result of the frames and the sequential nature of manga. The fact that the narrative has to be read in an exact order means that it is governed by the logic of time. The frames are similar to sentences in a novel. For instance, in Figure 53, the three narrow frames which follow the images act as an ellipsis, suggesting ‘and so on’. The narrow frames thus echo the three dots of an ellipsis.



Figure 53: The narrow frames suggest an ellipsis (Kishimoto 2007: 58)

Of course, it is possible for readers to skip frames but the full effect and meaning of the narrative depends on sequential reading. This treatment of the visual and written modes can be attributed to the nature of the Japanese writing system. The Japanese writing system, in particular the Kanji (Chinese) characters, are “basically images, however stylised” (Martinec 2003: 66) so treating writing as a visual entity is not entirely a new concept. However, the

fact that Kanji characters are writing means that they are also regarded as such. This means that when writing Kanji, a 'visual' entity is handled like as a written entity – it becomes governed by the logic of time.

Writing is, after all, as many theorists have noted, an image (Kress and van Leeuwen 1996; McCloud 1994; Eisner 1985). According to Kress and van Leeuwen “in alphabetic writing the image of the object represented has come, over time, to stand at first for the object, then for the abbreviation of the name of the object and eventually for its initial letter” (1996: 19). Consequently, McCloud describes writing as “abstract icons” (1994: 24). The extensive abstraction in alphabetic writing means that its connection to images has long been overlooked and it has over time evolved a different logic to that of writing. However, as we move from the era of page to that of screen, it is again becoming clear that writing is a visual entity and it is increasingly being treated as such (Jewitt 2004). What transpires from this analysis is that the affordances of the visual and the written modes are not strictly governed by the logic of space and time respectively. Rather the affordances of the modes are governed by their 'functional specialisation'. That is, individual users of a sign should decide which mode is best in representing the characteristics of a particular knowledge and whether that mode is the best in capturing the attention of the audience (Kress et al 2001, Kress 2003).

5.4 The influence of social and cultural practices on manga conventions

Manga is a comic genre but it is evident from this study that conventions employed in manga are distinctively different from that of Western comics. The influence of social and cultural practices on the conventions of a genre emerges as an important factor for the differences. Luke points out that “many educational descriptions of 'how texts work' tend to separate

analytically ideology from function” (1996: 318). In other words, while students are taught the code and conventions of a genre, they are not shown how the rules function as social strategies for instilling ideologies. By viewing genre as social practices situated in the context of situation and context of culture from which the genre emerges, it foregrounds the social constructedness of texts and genres.

This study has already mentioned the right-to-left reading direction in Japanese texts as a key difference between Western comics and manga. Another difference is the greater use of interpersonally- oriented resources in manga. McCloud (1994; 2006) has already noted that compared to Western comics, in manga there is a higher level of engagement or reader participation and a greater emphasis is placed on pacing out the narrative to create a sense of being there. These points have certainly surfaced in this analysis. The difference in the narrative approach can be attributed in part to the influence of film on manga but also to the social and cultural context of Japan. Unlike Western comics, which have origins in print and caricature drawings (Sabin 1996), manga draws its inspiration from film. The semiotic resources in manga therefore, often mimic that of film conventions, for example, point of view, camera distance and angle. These resources provide the reader a level of engagement similar to that found in film.

Another reason for the high level of participation in manga is due to the social and cultural context of Japan. In his comparative analysis of Japanese and English recipes, Martinec discovered that Japanese recipes tend to be “more elaborate in the extent to which they engage the reader/viewer, in the degree of detail with which they represent the portrayed action, and in the explicitness of marking the procedures’ stages” (2003: 43). He argues that

this is a result of the socio-cultural context of Japan. According to Martinec, Japan is a country where status differentiation is “finely graded”.

It appears that the Japanese find it rather difficult to interact without knowing each other’s social status, having to decide not only on one of several speech levels that they should adopt, but also on a host of other, non-verbal actions ranging from facial expressions to seating arrangements, all of which are rather strictly codified (2003: 61).

So, in a business setting, the client is always treated with respect and every effort is made to satisfy his/her needs. This kind of relationship also extends to the relationship between a producer and a consumer of text. This may explain why considerable effort is made to draw the reader into the narrative in manga. The high level of engagement acknowledges the presence of the reader and makes the reading experience more entertaining.

Martinec (2003) notes the degree of empathy in Japanese culture as an additional factor for the high level of engagement. Citing from Lebra (1976), Martinec writes that “[f]or the Japanese, empathy [*omoiyari*] ranks high among the virtues considered indispensable for one to be really human, morally mature, and deserving of respect” so a concerted effort is made to accommodate for the other’s needs (2003: 61). Perhaps this explains the extensive use of resources which highlight interpersonal relations such as close-ups, facial expressions and other emotionally expressive effects. These resources draw empathy from the reader.

It is thus clear that the social and cultural practices of a society greatly influence the conventions of a genre. This extends to the ways in which modes of communication are employed. The functional specialisation of modes derives from their affordances and “by repeated uses in a culture, or by the interested use of the individual sign-maker / designer” (Kress 2003: 46). This means that modes and their specialisations are socially oriented and

the manner in which they are employed may differ from culture to culture. In the West, logocentrism has resulted in the written mode being well developed as a communicational resource while other modes such as images have been largely neglected until recent years. This means that in Western texts, writing tended to dominate. From Martinec's research, he found that there was "greater communicative load of the visual mode in Japanese culture as compared with English, and, perhaps generally, Western culture" (2003: 65). He attributes this phenomenon to the stronger emphasis on face-to-face relationship in Japan and the nature of their writing system.

the Japanese writing system, and the pictographic and ideographic characters imported from China (*kanji*) in particular, is certainly a factor in the greater use of images as well...The prestige that the Japanese attach to teaching and learning *kanji* is surely a powerful aid in establishing a visual awareness very early in the school age (Martinec 2003: 66).

This indicates the fact that culture plays an important role in the development of modes and their semiotic potentials. As Kress suggests "[a] culture can work with or against affordances, for reasons that lie with concerns other than representation (Kress 2003: 46). In the West, the written mode is privileged at the expense of other modes of meaning because of the belief that it is the "instrument of cultural and scientific progress" (Cope and Kalantzis 2000: 217). According to Kress and van Leeuwen, the written mode is so weight with the concept of literacy that "the move towards a new literacy, based on images and visual design" is "seen as a threat, a sign of the decline of culture" (1996: 15). Of course, this struggle with modes of representation in literacy is actually a struggle over power and capital (Luke 1996). Nevertheless, such concepts have hindered the potential for images to grow as a semiotic system.

In sum, this study reveals that social and cultural factors play an important role in shaping genre conventions. These include influences on the functional specialisation and the functional load in a text. Taking into account social and cultural factors when analysing a text provides a better understanding of how and why texts work the way they do. Moreover, it points to the fact that conventions are social and cultural resources employed by individuals to produce texts for certain purposes. This in turn directs attention to the social constructedness of genres and the fact that they are products of design. This notion has implications for using a metalanguage of manga in interrogating other visual narratives.

5.5 The possible implications of using a metalanguage of manga in interrogating other visual narratives

The concept of design and the sign as motivated promotes a way of thinking which makes it possible to “harness students’ resources” (Archer 2006b, Archer 2004) in productive ways. Past literacy theories are characterised by ‘use’. A theory of use is governed by “a stable system with stable elements” (Kress 2003: 40). A stable system encourages standardised forms of meaning. This is best achieved through monomodality. In a theory of use, individuals are seen as users of the system. This means that “creativity is rare, it is special and exceptional” (Kress 2003: 40). The notion of design and the sign as motivated, however, is based on a social theory of semiotics where meaning is a result of work. According to Kress, “[w]ork always changes those who do the work, and it changes that which is worked on”. This means that “*creativity* is ordinary, normal; it is the everyday process of semiotic work as making meaning” (2003: 39, 40). Moreover, because the sign is motivated, a social theory of semiotics recognises that meaning from texts is always an approximation (Kress 2003). Individuals make hypothesis on how to interpret a sign based on previous encounters.

This suggests the possibility for individuals to negotiate meanings and make hypotheses about new genres and discourses based on what they already know. A metalanguage of analysis which is able to identify and discuss different forms of meaning can assist in the transformation process by creating a dialogue between old and new genres, old and new discourses. It can help students to recontextualise meanings, and apply what they know “in relation to other ways of knowing” (Thesen 2001: 143). In terms of this study, this means that it is possible to use a metalanguage of manga to examine other visual narratives. The following section proposes a case where a metalanguage of manga may be used to interrogate storyboarding.

5.6 Using a metalanguage of manga to examine storyboarding

A metalanguage of manga may be used to interrogate storyboarding by using the metalanguage to highlight the similarities and differences between manga and storyboarding conventions. The first necessary step is to contextualise the two genres.

Manga is a popular cultural text. It is a genre of comics therefore its function is to provide entertainment. This means that the semiotic resources employed are ‘embedded’ so that they provide readers a pleasant reading experience. The proposed metalanguage of analysis could be utilised to help students to recognise the resources at work, how they are employed and what ideologies they may evoke. This process would follow one similar to how the data in this study has been analysed.

The storyboard can be described as a sequence of shots sketched on paper and annotated with production directions. They are used for production purposes in the film and television

industry. In general, a storyboard “allow[s] a filmmaker to previsualize his [sic] ideas and refine them in the same way a writer develops ideas through successive drafts” (Katz 1991: 24). By having the shots planned out in advance, this process helps to save time and money at the production stage. Moreover, storyboards help to ensure a flow in the narrative. This is particularly important in a complex sequence of shots such as stunt scenes. A storyboard also “serve[s] as the clearest language to communicate ideas to the entire production team” (Katz 1991: 24). This process ensures that the crew (for example the camera or the lighting team) knows exactly what is expected from them. A storyboard is therefore a communicational tool, which directs people to perform specific functions. This means that resources that are used to film a shot need to be made explicit. According to Katz, a storyboard should convey two basic kinds of information: “a description of the physical environment of the sequence (set design/location) and a description of the special quality of a sequence (staging, camera angles, lens and the movement of any elements in a shot)” (1991: 44-45). Set design or location is usually communicated through the images themselves. Special qualities such as camera angles and movements are annotated outside the framed image. Figure 54 provides an example of the basic elements of a storyboard.

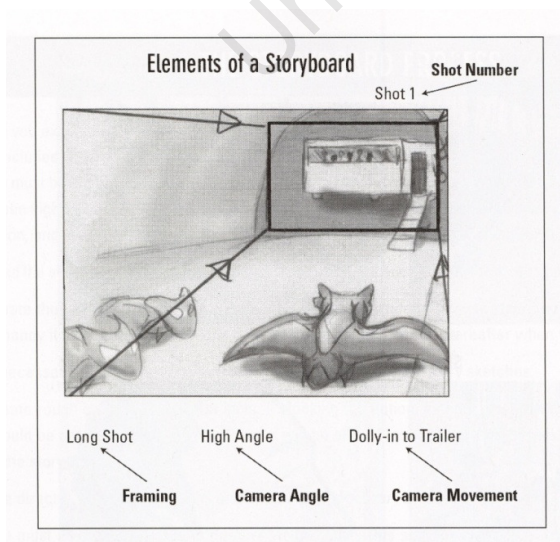


Figure 54: Elements of a storyboard (Tumminello 2005: 5)

From this description of the storyboard, it becomes clear that manga and storyboards are alike in many ways. Both are visual narratives, except the purpose of manga is to entertain so the semiotic resources are embedded while the purpose of a storyboard is to direct so the semiotic resources are made explicit. In reorganising the resources in manga, the intertextual elements in the two genres become even more evident. To demonstrate this point, images from Naruto and have been reorganised so that they mimic a storyboard. The images now depict diagrammatic arrows showing camera movement, the camera distance and angle are annotated and speech frames have been erased from the images. Dialogue is added as annotations (see Figure 55).

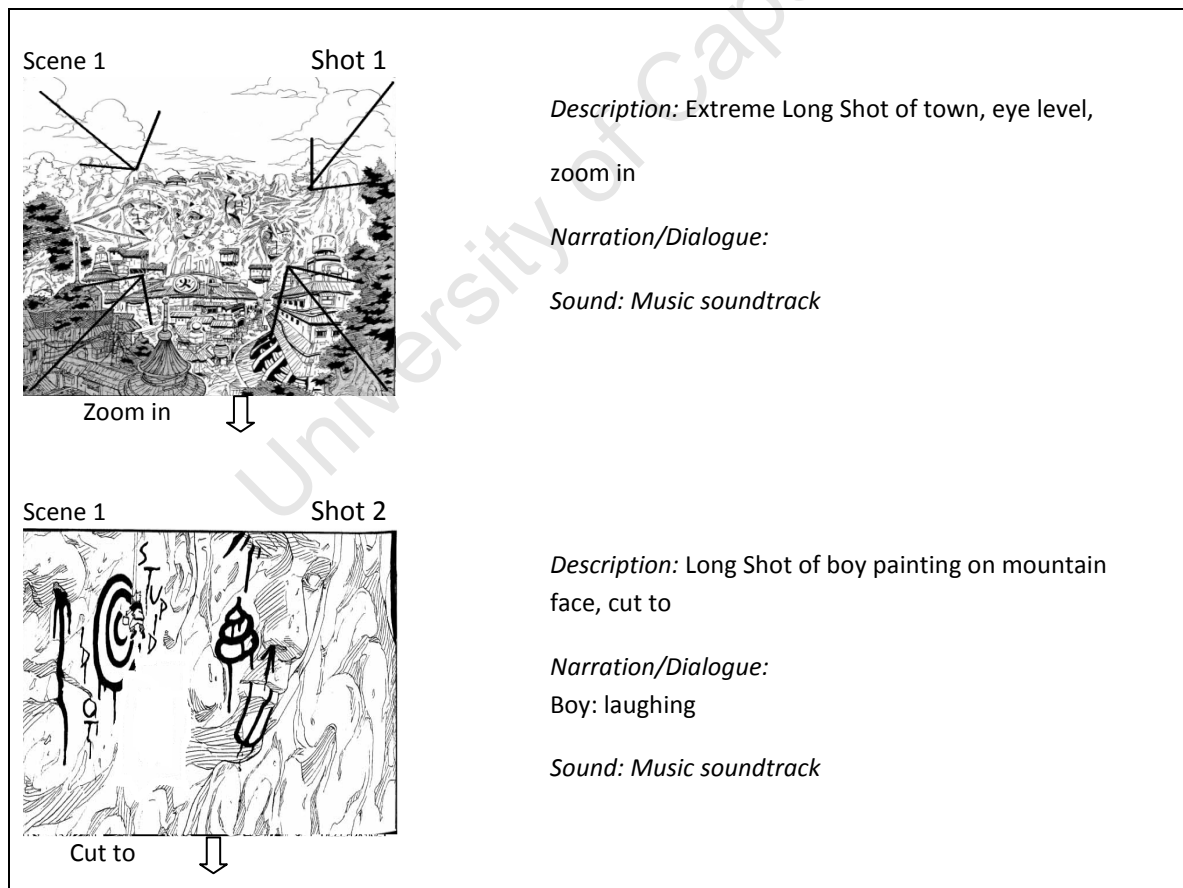


Figure 55: Transforming manga into a storyboard

According to Kress, “Design asks, ‘what is needed now, in this one situation, with the configuration of purposes, aims, audience, *and with these resources*, and given *my* interests in this situation” (2003: 49). Given that students already have knowledge of how conventions of a genre work, through the concept of design, they can transfer this knowledge into understanding another genre. In this case, an understanding of how manga narratives work can help students to construct a storyboard narrative. There are, of course, limitations. Some images in manga will not necessarily work in film. This is when the metalanguage of analysis is crucial in negotiating the limitations. What is the difference between manga and film? Why does the image work in manga but not in film? The metalanguage can thus assist students in understanding and interpreting difference between texts.

5.7 Final comments

This study has taken a multimodal social semiotic approach to analysing a manga narrative. In analysing the narrative, it has demonstrated that different semiotic resources employed in manga perform distinct narrative functions and it is important that the resources are read in relation to each other.

It has emerged that conventions of a genre are grounded in the social and cultural contexts from which it emerges. It is important to take these contexts into account as they affect the interpretation of texts. For example, the reading direction of Japanese culture greatly influences the reading direction in manga and the text composition as a whole. In highlighting the social constructedness of genres, this study foregrounded the notion of genre as ‘designs’ and semiotic resources as design resources. This study has argued that a

metalanguage which can discuss different forms of meaning can also assist individuals to see the similarities between genres by highlighting the use of conventions.

In examining the data, a metalanguage for manga was proposed. The metalanguage is predominantly based on Kress and van Leeuwen's (1996) 'grammar for visual design' as well as Matthiessen's (2007) notion of rhetorical relations. Included in the metalanguage are representational resources such as facial expression, body posture and gesture.

A major contribution of the study is that it extends understanding of the nature of sequential visual narratives. It contributes to a better understanding of the affordances of various semiotic resources and how they may be employed in narratives of various kinds.

University of Cape Town

References

- Alvermann, D.E. and Heron, A.H. 2001. Literacy identity work: playing to learn with popular media. In *Journal of Adolescent & Adult literacy*. 45. 118-122.
- Archer, A. 2008. Cultural studies meets academic literacies: exploring students' resources through symbolic objects. In *Teaching in Higher Education*. 13 (4). 383-394.
- Archer, A. 2006a. A multimodal approach to academic 'literacies': Problematizing the visual/verbal divide. In *Language and Education*. 20 (6). 449-462.
- Archer, A. 2006b. Opening up spaces through symbolic objects: Harnessing students' resources in developing academic literacy practices in engineering. In *English Studies in Africa*. 49 (1). 189-206.
- Archer, A. 2004. *Access to academic practices in an engineering curriculum: drawing on student' representational resources through a multimodal pedagogy*. PhD thesis University of Cape Town, Cape Town.
- Baldry, A.P. and Thibault, P.J. 2006. *Multimodal transcription and text analysis. A multimedia toolkit and coursebook*. London and New York: Equinox.
- Barthes, R. 1967. *Elements of semiology*. London: Cape.
- Bhatia, V. 2004. *Worlds of written discourse*. London: Continuum.
- Branigan, E. 1992. *Narrative comprehension and film*. London: Routledge.
- Brenner, R.E. 2007. *Understanding Manga and Anime*. Westport, CT: Libraries Unlimited/Greenwood.
- Chandler, D. 2007. *Semiotics: the basics*. London; New York : Routledge.
- Chatman, S. 1978. *Story and discourse: narrative structure in fiction and film*. Ithaca: London Cornell University Press.
- Cohen, L., Manion, L. and Morrison, K. 2007. *Research methods in education*. London, New York: Routledge.
- Cope, B. and Kalantzis M. 2006. From Literacy to 'Multiliteracies'. In *English Studies in Africa*: 49 (1). 23-45.
- Cope, B. and Kalantzis, M. (Eds). 2000. *Multiliteracies: literacy learning and the design of social futures*. London: Routledge.
- Cope, B. and Kalantzis, M (Eds). 1993. *The Powers of literacy: a genre approach to teaching writing*. London: Falmer Press.

- Douglas, M. 2005. *Purity and danger : an analysis of concept of pollution and taboo*. London: New York: Routledge.
- Eisner, W. 1996. *Graphic Storytelling*. Tamarac, Florida: Poorhouse Press.
- Eisner, W. 1985. *Comics and Sequential Art*. Tamarac, Fla: Poorhouse Press.
- Foucault, M. 1995. *Power/knowledge: selected interviews and other writings*. Gordon, C. (Ed). New York: Harvester Wheatsheaf.
- Fairclough, N. 1995. *Critical discourse analysis: the critical study of language*. London: Longman.
- Fairclough, N. 1992. *Discourse and social change*. Cambridge: Polity Press.
- Gee, J.P. 2003. *What video games have to teach us about learning and literacy*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
- Gee, J. 1999. *An introduction to discourse analysis: theory and method*. London, New York: Routledge.
- Groensteen, T. 2000. Why are comics still in search of cultural legitimization? In *Comics and culture: Analytical and theoretical approaches to comics*. Magnussen, A. and Christiansen, H. (Eds.). Copenhagen, DK: Museum Tusulanum Press. 29-42.
- Hall, S. 1997. *Representation: cultural representations and signifying practices*. London: Sage in association with the Open University.
- Halliday, M.A.K. 1985. *An Introduction to Functional Grammar*. London: Arnold.
- Halliday, M.A.K. and Hasan, R. 1985. *Language, Context, and Text: Aspects of Language in a Social-Semiotic perspective*. Belmont, Vict: Deakin University.
- Harvey, R. 1996. *The Art of the Comic Book: An Aesthetic History*. Jackson: University Press of Mississippi.
- Hodge, R. and Kress, G. 1988. *Social Semiotics*. Ithaca, N.Y: Cornell University Press.
- Ito, K. 2005. A history of manga in the context of Japanese culture and society. In *The Journal of Popular Culture*. 38(3). 456-475.
- Jewitt, C. 2004. Multimodality and new communication technologies. In *Discourse and Technology: Multimodal Discourse Analysis*. Levine, P. and Scollon, S (Eds.). Washington D.C.: Georgetown University Press. 84-195.
- Jewitt, C. and Oyama, R. 2001. Visual meaning: A social semiotic approach. In *Handbook of visual analysis*. van Leeuwen, T. and Jewitt, C. (Eds.). 134–156.

Katz, S. 1991. *Film Directing Shot by Shot: Visualising from Concepts to Screen*. California: Michael Wiese Productions. 23-82.

Kinsella, S. 2000. *Adult manga: culture and power in contemporary Japanese society*. Richmond, Surrey: Curzon.

Kress, G. 2003. *Literacy in the new media age*. London: Routledge.

Kress, G. 2000. Design and Transformation: New theories of meaning. In *Multiliteracies: literacy learning and the design of social futures*. Cope, B. and Kalantzis, M. (Eds). London: Routledge.

Kress, G. 1998. Visual and verbal modes of representation in electronically mediated communication. In *Page to Screen*. Snyder, I (Ed.). London: Routledge.

Kress, G., 1993. Genre as social process. In *The Powers of Literacy: A Genre Approach to Teaching Writing*. Cope, B. and Kalantzis, M. (Eds). Falmer Press, London. 22–37.

Kress, G., Ogborn, C., Jewitt, C. and Tsatsarelis, C. 2001. *Multimodal teaching and learning: the rhetorics of the science classroom*. London: Continuum.

Kress, G. and Threadgold, T. 1988. Towards a social Theory of Genre. In *Southern Review*. 21(3). 215-43.

Kress, G. and van Leeuwen, T. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.

Kress, G. and van Leeuwen, T. 1996. *Reading images: the grammar of visual design*. London: Routledge.

Labov, W. 1972. *Language in the inner city: studies in Black English vernacular*. Oxford: Blackwell.

Lim, V.F. 2007. The visual semantics stratum: making meaning in sequential images. In *New directions in the analysis of multimodal discourse*. Royce, T.D. and Bowcher, W.L. (Eds). Mahwah, N.J.: L. Erlbaum Associates. 195-214.

Luke, A. 1996. Genres of Power? Literacy Education and the Production of Capital. In *Literacy in Society*. Hasan, R. and Williams, G. (Eds). London New York: Longman. 308-338.

Martin, J.R. and Rose, D. 2003. *Working with discourse: meaning beyond the clause*. London, New York: Continuum

Martinec, R. 2003. The Social Semiotic of Text and Image in Japanese and English Software Manuals and Other Procedures. In *Critical Social Semiotics* (Special Issue). Van Leeuwen, T. and Caldas-Coulthard, C. 13 (1). 43-69.

Martinec and Salway. 2005. A System for Image-Text Relations in New (and Old) Media. In *Visual Communication*. 4(3). 337-371.

Matthiessen, C.M.I.M. 2007. The Multimodal Page: A Systemic Functional Exploration. In *New directions in the analysis of multimodal discourse*. Royce, T.D. and Bowcher, W.L. (Eds). Mahwah, N.J.: L. Erlbaum Associates. 1-62.

McCloud, S. 2006. *Making Comics: storytelling secrets of comics, manga and graphic novels*. New York, London, Toronto, Sydney: Harper.

McCloud, S. 1994. *Understanding Comics: the invisible art*. New York: HarperPerennia.

Neale, S. 2000. *Genre and Hollywood*. London, New York: Routledge.

New London Group. 2000. A Pedagogy of Multiliteracies: Designing Social Futures. In *Multiliteracies. Literacy Learning and the Design of Social Futures*. Cope, B. and Kalantzis, M. (Eds). London and New York: Routledge.

Norton, B. and Vanderheyden, K. 2005. Comic book culture and second language learners. In *Critical Pedagogies and Language Learning*. Norton, B. and Toohy, K. (Eds). Cambridge, England : Cambridge University Press. 201-222.

Prince, G. 1988. *A dictionary of narratology*. Aldershot: Scolar.

Rubinstein-Ávila, E. and Schwartz, A. 2006. Understanding the Manga Hype: Uncovering the Multimodality of Comic-Book Literacies. In *Journal of Adolescent & Adult Literacy*. 50(1). 40-49.

Ryan, M. (Ed). 2004. *Narrative across media: the language of storytelling*. Lincoln: University of Nebraska Press.

Sabin, R. 1996. *Comics, commix and graphic novels*. London: Phaidon.

Sabin, R. 1993. *Adult comics: an introduction*. London: Routledge.

Saussure, F. 1966. *Course in general linguistics*. Bally, C. and Sechehaye, A. New York: McGraw-Hill.

Stenglin, M.K. 2009. Space odyssey: towards a social semiotic model of three-dimensional space. In *Visual Communication*. 8 (35). 35-64.

Thesen, L. 2001. Modes, Literacies and Power: A University Case Study. In *Language and Education*. 15 (2 & 3). 132-145.

Toolan, M. 1988. *Narrative: a critical linguistic introduction*. London, New York: Routledge

Tumminello, W. 2005. *Exploring Storyboarding*. Australia: Thomson/Delmar Learning.

Unsworth, L. 2007. Multiliteracies and multimodal text analysis in classroom work with children's literature. In *New Directions in the Analysis of Multimodal Discourse*. Royce, T.D. and Bowcher, W.L. (Eds). Mahwah, N.J.: L. Erlbaum Associates. 331-360.

Unsworth, L. 2006. Towards a Metalanguage for Multiliteracies Education: Describing the Meaning-Making Resources of Language-Image Interaction. In *English Teaching: Practice and Critique*. 5 (1). 55-76.

van Leeuwen, T. 2005. *Introducing Social Semiotics*. London, New York: Routledge.

van Leeuwen, T. 2004. Ten reasons Why Linguists Should Pay Attention to Visual Communication. In *Discourse and technology: multimodal discourse analysis*. LeVine, P. and Scollon, T. (Eds). Washington D.C.: Georgetown University Press.

Wisker, G. 2008. *The postgraduate research handbook: succeed with your MA, MPhil, EdD and PhD*. Basingstoke, Hampshire, New York: Palgrave Macmillan.

Websites:

Eason, G. 2007. *Shakespeare gets comic treatment*.
http://news.bbc.co.uk/2/hi/uk_news/education/6647927.stm Accessed 4 February 2008.

<http://en.wikipedia.org/wiki/Manga> Accessed 1st August 2009.

Rommens, A. 2000. Manga story-telling/showing. *Image and Narrative: Online Magazine of the Visual Narrative*. <http://www.imageandnarrative.be/narratology/aarnoudrommens.htm>
Accessed 8 May 2008.

Figures:

Kishimoto, M. 2007. *Naruto*. Volume 1. San Francisco: VIZ Media.

Kishimoto, M. *Naruto*. Jump Comics.
www.narutofan.com Accessed 18 May 2008.

Cover page image:

www.animewallpapers.com Accessed 2 September 2009