

AN ABSTRACT OF THE DISSERTATION OF

Pallabi Chattopadhyay, submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Visual Linguistics, presented on 14 November 2025, titled *Cross-Linguistic Variations in Visual Narratives: A Study of Selected Picture Stories*.

Before the External Examiners: Prof. Dr. Arun Gupto and Prof. Dr. Kaiser Hamidul Haq

Approved by:

Dr. Uddhab Pyakurel

Dean of the School of Arts

Abstract Approved:

.....

Signature of Supervisor

Prof. Dr. Abhi Subedi

.....

Signature of Co-Supervisor

Prof. Dr. Sagar Raj Sharma

Abstract

Art is a form of expression born from human imagination and creativity that engages our visual perception in ways that allow each person to interpret it uniquely. Throughout history and across cultures, art has served as a fundamental mode of human expression, allowing communities to convey stories, identities, and shared experiences. Over time, art evolved from realistic expressions to abstract and surrealistic forms, eventually giving rise to visual narratives (comics, manga, and other forms) as unique blends of storytelling and visual art.

In this context, this dissertation explores the cross-linguistic dimensions of visual narratives, focusing on the stimuli—Bengali comics, *Hada Bhoda* and the Japanese manga, *Doraemon*, alongside *Patua* and *Emaki* scroll paintings from the two cultures of West Bengal (India) and Japan, respectively. A qualitative research approach has been used, employing content analysis methodology to investigate the grammar, structure and functions, along with temporal and communicative aspects of Linguistics.

Through the thematic coding provided by the proposed theoretical framework, the research emphasises how visual narratives share universal connections while also having distinct differences. It analyses the grammatical, structural and functional elements to reveal their roles in cross-cultural communication. Furthermore, it evaluates synchronic and diachronic variations, tracing the evolution of the selected picture stories from traditional scrolls to modern-day Bengali comics and Japanese manga. The study also applies speech act theory through communicative acts, facial expressions, ideophones and drawing styles, and the findings explore how the culturally nuanced nature of visual narratives offers broader comprehension into multimodality and multisensory aspects of visual storytelling. As a result, the study observes new dimensions and ideas in the field of Visual Linguistics and opens pathways for further research into the world of visual storytelling.

APPROVAL OF THE DISSERTATION

Dissertation by Pallabi Chattopadhyay submitted to the School of Arts Research Committee, Kathmandu University, on 10 July 2025.

Title: Cross-Linguistic Variations in Visual Narratives: A Study of Selected Picture Stories

APPROVED

.....

Ms. Ekku Maya Pun

Dean of the School of Arts
Chairperson, Research Committee

10 July 2025

.....

Prof. Dr. Sagar Raj Sharma

Research Committee Member

10 July 2025

.....

Prof. Dr. Nirmala Mani Adhikary

Research Committee Member

10 July 2025

.....

Prof. Dr. Laya Prasad Uprety

Research Committee Member

10 July 2025

.....

Dr. Savitree Thapa Gurung

Research Committee Member

10 July 2025

Dedication

To Papa and Mamoni, who have been my constant support and endless inspiration. Your love, guidance, and belief in me have made this journey possible. Thank you for inspiring a sense of purpose in me and also for letting me dream...

...And, to my loving niece, Oishika, whose love for art reminds me of the magic in imagination and creativity! May this work inspire you to follow your passions and fill your world with the wonders of colours and your unique vision...

To Papa, whose dreams became the foundation of my own. This PhD is not just mine; it is yours. Every step of this journey was guided by your faith in me, your quiet strength, and the values that you passed on to me. You left before my final defence, but your presence is always there in my heart. Wherever you are, I hope you know that this achievement is for you and because of you.

I had written the first half of the dedication before my final presentation, never knowing that you would leave me before I could finish it. Now, as I complete this work, every word carries your memory with it. This is, and will always be, for you.

Declaration

I hereby declare that this dissertation has not been submitted for candidature for any other degree. I understand that my dissertation will become a part of the permanent collection of Kathmandu University Library. My signature below authorises the release of my dissertation to any reader upon request.

Signature of Student:

Date:

.....

Acknowledgments

I would first like to express my deepest gratitude to my supervisors, Prof. Dr. Abhi Subedi and Prof. Dr. Sagar Raj Sharma, for believing in the potential of this research project as a significant contribution to the field of Linguistics. I also thank them for their expertise, guidance, and support throughout this Ph.D. journey. Their constructive feedback and encouragement have shaped this dissertation and my academic growth in countless ways.

I am also deeply grateful to Kathmandu University, School of Arts, the Department of Languages and Mass Communication, and all the professors, especially to the Head of the Department, Prof. Dr. Nirmala Mani Adhikari, for providing an enriching academic environment with resources; special thanks to the Dean of the School of Arts, Dr. Uddhab Pyakurel, for his support in numerous ways.

There are no words that can fully capture the depth of my gratitude to my dear sister, Pritha, and brother-in-law, Rajdip. Their unwavering support throughout this entire Ph.D. journey has been nothing short of extraordinary. From their generous financial support to their emotional encouragement, they have stood by me like pillars. A simple thank you could never be enough to express my appreciation for all they have done for me to reach this point.

I am also deeply thankful to my parents, who provided me with constant moral support, financial help, and their unfailing tolerance.

Finally, my thanks go to all my well-wishers and friends who have stood by me through thick and thin. Especially, I would like to express my heartfelt gratitude to my dear friends, Lahari and Roshan, for their steadfast support throughout this PhD journey. Roshan, in particular, has been a true friend away from home. From thoughtfully cooking for me to sharing academic discussions, he was always there with constant support and understanding.

Table of Contents

Introduction: Visual Linguistics and Visual Narrative Structures	1
1.1 Background of the Study	1
1.1.1 Visual Linguistics: A Part of Linguistic Studies	7
1.1.2 Visual Linguistics and The Theory of Innateness	8
1.1.3 Cross-Cultural Perceptions in Visual Linguistics.....	10
1.2 Statement of the Problem	15
1.3 Objectives and Research Questions	21
1.3.1 Theoretical Foundation Guiding the Objectives for Cross-Linguistic Analysis	24
1.4 Significance and the Study	26
1.5 Key Operational Terms and Narrative Frameworks in Visual Narratives	29
1.6 Outline of Dissertation	33
Chapter 2.....	36
Literature Review: Conceptualising the Study	36
2.1 Methodological Review	36
2.2 Thematic Review.....	39
2.2.1 Cross-Linguistic Analysis.....	40
2.2.2 McCloud on Comics.....	42
2.2.3 Cohn’s Views on Visual Narratives	44
2.3 Theoretical Review	46
2.3.1 The Chomskian Hypothesis: The Innate Module of Language.....	48

2.3.2 Halliday’s Systemic Functional Linguistics: Its Relation to Chomskian Theory and Visual Linguistics	51
2.3.3 Wittgenstein and Pinker in Visual Linguistics	56
2.3.4 Diachronic and Synchronic Linguistics in Visual Narratives	61
2.3.5 Speech Act Theory in Visual Narratives	63
2.4 Review of Previous Research.....	66
2.5 Research Gap: The Unexplored Dimensions	69
Chapter 3.....	73
Research Methodology	73
3.1 Research Philosophy	73
3.2 Research Approach	76
3.3 Theoretical Framework	78
3.4 Conceptual Design	81
3.5 Research Method.....	83
3.6 Sampling Design	86
3.7 Method of Data Collection.....	92
3.8 Method of Data Analysis and Interpretation	93
3.9 Limitations and Delimitations.....	95
3.9.1 Limitations.....	96
3.9.2 Delimitations	97
Chapter 4.....	99

Findings (Objective 1): A Cross-Linguistic Analysis of Visual Narratives through the Grammar, Structure and Social Contexts.....	99
4.1 TGL and SFL in Visual Narratives	99
4.2 Structure and Grammar of Visual Narratives (<i>Hada Bhoda</i> and <i>Doraemon</i>).....	99
4.3 Functional Aspects of Visual Narratives- Society in Miniature	108
4.3.1 Family Dynamics	108
4.3.2 Friendly Exchanges	114
4.3.3 School Conversations	117
4.3.4 Sportsmanship	121
4.4.5 Culinary Culture	123
4.4 Discussion	127
Chapter 5.....	131
Findings (Objective 2): A Cross-Linguistic Analysis of Visual Narratives through their Temporal Aspects	131
5.1 Temporal Aspects in Visual Narratives	131
5.2 Order of Panels in Scroll Paintings	135
5.3 Social Contexts.....	142
5.4 Discussion	149
Chapter 6.....	154
Findings (Objective 3): A Cross-Linguistic Analysis of Visual Narratives through Speech Act Theory	154
6.1 Speech Acts in Visual Narratives.....	154
6.2 Communicative Acts.....	154

6.3 Facial Expressions.....	161
6.4 Ideophones	165
6.5 Drawing Styles and Multimodality	173
6.6 Discussion	176
Chapter 7.....	181
Final Insights: Cross-Cultural Visual Narratives and their Cross-Linguistic Variations	181
Where the Journey Ends, Another Begins... ..	181
7.1 Summary	181
7.1.1 The Interplay of Culture in Visual Narratives	181
7.1.2 Final Reflections on Theoretical Framework	184
7.2 Findings in Congruency with the Objectives	185
7.3 Concluding Reflection and Implications of the Study	187
7.3.1 Theoretical Implications	189
7.3.2 Methodological Implications	189
7.3.3 Pedagogical Implications.....	189
7.4 Expanding Horizons and Final Thoughts.....	191
References.....	197

List of Figures

Figure 1. Initial Framework	25
Figure 2. Picture Story: The Captivating Book	53
Figure 3. Picture Story: The Last Apple	60
Figure 4. Theoretical Framework	80
Figure 5. Conceptual Design.....	82
Figure 6. Semantic Shifts Showing the Grammar and Structure of Bengali Comics and Japanese Manga	100
Figure 7. Semantic Shifts of Hada Bhoda and Doraemon	101
Figure 8. A Moment-to-Moment Transition.....	102
Figure 9. An Action-to-Action Transition	103
Figure 10. A Subject-to-Subject Transition.....	104
Figure 11. A Scene-to-Scene Transition.....	105
Figure 12. Aspect-to-Aspect Transition.....	106
Figure 13. A Non-Sequitur Transition	107
Figure 14. Debnath's (1962) Hada Bhoda Samagra (p. 61)	108
Figure 15. Debnath's (1962) Hada Bhoda Samagra (p. 490)	109
Figure 16. Debnath's (1962) Hada Bhoda Samagra (p. 392)	110
Figure 17. Fujio's Doraemon- A Selection of Touching Stories (p. 042)	111
Figure 18. Fujio's Doraemon- Volume 4 (p. 140).....	112
Figure 19. Fujio's Doraemon- Volume 2 (p. 28) (Read: Right to Left)	113
Figure 20. Debnath's (1962) Hada Bhoda Samagra (p. 287)	114
Figure 21. Debnath's (1962) Hada Bhoda Samagra (p. 462)	115
Figure 22. Debnath's (1962) Hada Bhoda Samagra (p. 462)	115
Figure 23. Fujio's (1974) Doraemon- A Selection of Touching Stories (p. 005).....	116
Figure 24. Fujio's (1974) Doraemon- Volume 4 (p. 110) (Read: Right to Left).....	116

Figure 25. Debnath's (1962) Hada Bhoda Samagra (p. 123)	117
Figure 26. Debnath's (1962) Hada Bhoda Samagra (p. 125)	118
Figure 27. Fujio's (1974) Doraemon- Volume 3 (p. 3)	119
Figure 28. Debnath's (1962) Hada Bhoda Samagra (p. 50)	119
Figure 29. Fujio's (1974) Doraemon- Volume 1 (p. 53)	120
Figure 30. Fujio's (1974) Doraemon- Volume 4 (p. 120)	120
Figure 31. Debnath's (1962) Hada Bhoda Samagra (p. 16)	121
Figure 32. Debnath's (1962) Hada Bhoda Samagra (p. 18)	121
Figure 33. Fujio's (1974) Doraemon- Volume 5 (p. 52)	122
Figure 34. Fujio's (1974) Doraemon- Volume 1 (p. 133) (Read: Right to Left).....	123
Figure 35. Debnath's (1962) Hada Bhoda Samagra (p. 71)	123
Figure 36. Debnath's (1962) Hada Bhoda Samagra (p. 71)	124
Figure 37. Fujio's (1974) Doraemon- Volume 5 (p. 22)	125
Figure 38. Debnath's (1962) Hada Bhoda Samagra (p. 71)	125
Figure 39. Fujio's (1974) Doraemon- Volume 5 (p. 22)	126
Figure 40. Fujio's (1974) Doraemon- Volume 4 (p. 57)	126
Figure 41. Ranjit Chitrakar, giving me Information about Patua Scroll Paintings.....	132
Figure 42. Patua Scroll Painting (Achronological Ordering)	135
Figure 43. Patua Scroll Painting (Single Panel).....	136
Figure 44. Patua Scroll Painting (Twin Panel)	136
Figure 45. Patua Scroll Painting (Chronological Ordering)	137
Figure 46. Patua Scroll Painting (Multiple Panels)	137
Figure 47. Patua Scroll Painting (Multiple Panels- Z-Path)	138
Figure 48. Patua Scroll Painting (Multiple Panels, Left-to-Right)	139
Figure 49. Emaki Scroll Painting- Natural Gutters and Text.....	140
Figure 50. Emaki Scroll Painting- Alternation between Text and Painting	140

Figure 51. Emaki Scroll Painting- Successive Painted Scenes without Text	141
Figure 52. Emaki Scroll Painting- Successive Painted Scenes with Two Sections of Text	141
Figure 53. The Ramayana- Old Patua Scroll Painting (Left); New Patua Scroll Painting (Right).....	142
Figure 54. Debnath's (1962) Hada Bhoda Samagra (p. 423)	143
Figure 55. Patua Scroll Painting (Machher Biye).....	143
Figure 56. Debnath's (1962) Hada Bhoda Samagra (p. 214)	144
Figure 57. Patua Scroll Painting (Marriage and Society)	145
Figure 58. Debnath's (1962) Hada Bhoda Samagra (p. 308)	146
Figure 59. Genji Monogatari Emaki- Yadorigi Gi Chapter and Text Calligraphy	146
Figure 60. Fujio's (1974) Doraemon- Volume 1 (p. 13)	147
Figure 61. Emaki scroll painting- Chōjū-giga	148
Figure 62. Fujio's (1974) Doraemon- Volume 1 (p. 27)	148
Figure 63. Debnath's (1974) Hada Bhoda Samagra (p. 252)	156
Figure 64. Debnath's (1962) Hada Bhoda Samagra (p. 328)	157
Figure 65. Fujio's (1974) Doraemon- Volume 2 (p. 77)	159
Figure 66. Fujio's (1974) Doraemon- Volume 4 (p. 103)	160
Figure 67. The Facial Expressions found in Debnath's (1962) Hada Bhoda Samagra	162
Figure 68. The Facial Expressions found in Fujio's (1974) Doraemon series	163
Figure 69. Debnath's (1962) Hada Bhoda Samagra (p. 38)	166
Figure 70. Debnath's (1962) Hada Bhoda Samagra (p. 415)	166
Figure 71. Debnath's (1962) Hada Bhoda Samagra (p. 378)	167
Figure 72. Debnath's (1962) Hada Bhoda Samagra (p. 68)	167
Figure 73. Debnath's (1962) Hada Bhoda Samagra (p. 31)	168
Figure 74. Fujio's (1974) Doraemon- Volume 4 (p. 130)	168
Figure 75. Fujio's (1974) Doraemon- A Selection of Touching Stories (p. 100).....	169

Figure 76. Fujio’s (1974) Doraemon- A Selection of Touching Stories (p. 103).....	169
Figure 77. Fujio’s (1974) Doraemon- Volume 1 (p. 114)	170
Figure 78. Fujio’s (1974) Doraemon- A Selection of Touching Stories (p. 046).....	170
Figure 79. Debnath’s (1962) Hada Bhoda Samagra (p. 15)-Left; Fujio’s (1974) Doraemon- A Selection of Touching Stories (p. 046)-Right	171
Figure 80. Fujio’s (1974) Doraemon- Volume 5 (p. 10)- Left; Debnath’s (1962) Hada Bhoda Samagra (p. 487)-Right.....	172
Figure 81. Debnath’s (1962) Hada Bhoda Samagra (p. 409)-Left; Fujio’s Doraemon- Volume 4 (p. 046)-Right.....	173
Figure 82. Debnath’s (1962) Hada Bhoda Samagra (p. 409)- Left; Fujio’s (1974) Doraemon- Volume 4 (p. 046)-Right.....	174
Figure 83. Similar Expressions	175

Chapter 1

Introduction: Visual Linguistics and Visual Narrative Structures

1.1 Background of the Study

‘Painting is poetry that is seen rather than felt, and poetry is painting that is felt rather than seen.’ — Leonardo da Vinci

Art is an expression used by human imaginations and creative abilities; it demands visual perceptions, thereby a unique comprehension takes place in each individual human brain. In fact, it is something that we visualise and perceive with the help of our brain and eyes. In this context, it can be said that dream and art are quite alike, as they are both visually interpreted yet are not the so-called ‘actual realities’. Thus, it seems that pleasure and pain in dreams are not real in their physical states; they exist in our brains mentally. Visual Linguistics should also be about the depths of the mind and the revealing of the so-called ‘false reality’. Visual Linguistics links the mind with the external world by means of visuals. It is to be clarified that visualisation has a direct link with human imagination, which is a unique property possessed only by human beings. The notion that art and dreams share a kinship in their detachment from objective reality makes it prominent that visual interpretation is not merely a cognitive act but also a deeply emotional and psychological experience. Visual Linguistics, then, must extend beyond structure, form, and semiotics; it should engage with the different dimensions of perceptions and the internalised stories we project onto the images we consume. This interpretative act is rarely neutral—it is shaped by memory, desire, fear, culture, and personal history. In both dreams and art, what we see is filtered through layers of unconscious thought, cultural frameworks, and emotional landscapes.

To engage with Visual Linguistics is to engage with a field that links between what is seen and what is understood, between the external stimulus and the internal narrative. Visuals are

never static; they live and breathe differently in each and every mind. A child from a rural region of West Bengal might look at a visual of a rice field and find comfort or familiarity, while a child from an urban landscape in Japan might perceive the same image as exotic or distant. The act of interpretation is steeped in this subjectivity. That's why it can be argued that visual texts can never be wholly universal.

Furthermore, the concept of 'false reality'—which exists vividly in the mind but not in the physical world—is crucial to this discussion. Dreams are real to the dreamers while they are being dreamt. A nightmare can induce actual fear responses like sweating or heart pounding, even though the perceived threat is not physically present. Similarly, a painting or a panel from a visual narrative can stir real emotions: grief, joy, or nostalgia. These responses may be intangible, but they are valid, real, and worth exploring. Visual Linguistics must take into account these ephemeral experiences because they form the very core of human engagement with visual media.

If language is often seen as a symbolic system through which reality is constructed and communicated, then visual language performs a parallel function. It constructs alternative realities that are no less valid simply because they do not conform to the physical world. What is seen, and how it is seen, has the power to shape internal belief systems, values, and emotional responses. The way an eye moves across a comic panel, the way a scroll painting is unrolled and narrated, the symbolism embedded in a facial expression—all these are tools of an emotional and cognitive choreography that guide meaning-making.

This leads to another critical dimension: temporality. Just as dreams unfold in fragmented sequences and jumps in time, so too do many visual narratives, especially those from non-Western traditions like the *Patua* scrolls or Japanese *Emakimono*. These narratives often defy the linear logic prized by Western visual conventions, presenting time as cyclical, fluid, or emotionally constructed. Visual Linguistics must be capable of addressing these alternative

models of time, where the sequencing of events is governed by mood, rhythm, or symbolism rather than strict chronology.

According to Locke (1689), the only things we perceive immediately are ideas. Locke's understanding of knowledge of the external world within its broader epistemology and theoretical philosophy requires probing beyond his epistemology and the depths of its accounts of perception, representation, and contents of thoughts. Therefore, the real difference between humans and other animals lies at the collective level. Unlike other animals, human beings rule the world because we can cooperate flexibly in large numbers. This is because human beings are the only organisms that live in a state of dual reality- a reality of ideas and conceptions, along with objective reality like other animals (Harari, 2015). Human beings possess this dual reality because of the cognitive revolution (c. 70,000 BCE, when Sapiens evolved imagination). As a result of this, humans can form stories or ideas, and eventually, they give birth to creation. Every inanimate object that is present around us physically was once created by a human being, which had previously originated in their idea. An artwork, much like a dream, is a creation or an imitation of the physical world (Plato, 1943). Unlike dreams, which disappear upon waking up, art can be preserved and maintained as a solid depiction of reality. Much like poetry, it captures and shares our comprehension of the world across generations. On a surface level, one can say that an artwork appears to be a visual story or narrative crafted with creative strokes to transmit ideas or messages from one human being to another. A visual interpretation acts as a bridge between the transient and the permanent. While our dreams are brief and fleeting, art and other visual narratives enable us to store and keep those disposable moments, giving them meaning and depth. It is a transformation of intangible emotions into the physical world, giving us ways to express and perceive our deepest ideas and expressions. Moreover, it can also be said that art goes beyond the boundaries of language, providing a universal form of communication. Colour, style, form, and composition allow artists to convey complicated

thoughts and trigger deep emotions that are difficult to express with words alone. Visual art creates meaning and comprehension through its own distinct language of symbols and aesthetics, just as poetry does through rhythm and metaphor.

Visual Linguistics should not only analyse images for what they show but also for what they awaken. It should study how visual forms mirror inner consciousness, how composition evokes subconscious archetypes, and how colour and spatial organisation can trigger memories or primal responses. Ultimately, recognising the alignment between dreams, art, and visual language invites scholars to question traditional binaries like real/unreal, true/false, and even consciousness/unconsciousness. In doing so, Visual Linguistics becomes not just a technical or analytical field, but a philosophical and psychological one, too, asking not just how visuals communicate, but why they matter so deeply to us as human beings. It is this deeper pursuit of meaning that gives Visual Linguistics its power, allowing it to delve into the folds of perception and uncover the intricate web of reality, imagination, and cultural identity.

Art is a medium that has been our companion since time immemorial, and it plays an especially important role in human lives; it is a way of expressing human feelings. From the time of cave paintings and rock paintings in the Paleolithic Age to the present time, where it has taken a commercial turn, art has always helped us express our emotions in different ways. Many civilisations like Ancient Egypt, Mesopotamia, India, China, Persia, Ancient Greece, Rome, Maya, and Olmec have some foundation and development in their own unique styles of art (Adhikari, 2022). Some of these styles were even transmitted to other cultures over time. Therefore, the most important aspect of art is that it preserves human cultures around the world. Ancient monuments, designing patterns, and even scripts add to such preservation. Adhikari states that these preservations help us to understand our past and help the new generation to perceive the roots of human civilisation.

Gradually, with time, artistic expression has evolved into many forms. From concrete and

realistic artworks, human beings have evolved to abstract and surrealistic forms of art. One such surrealistic form that is quite in vogue these days is cartoon drawings. According to Merriam-Webster Online Dictionary, the term ‘cartoons’ initially referred to an artist’s preparatory designs or paintings as models for the finished work. This shift or evolution is somewhat similar to the metamorphosis of Coca-Cola, which was first produced by a chemist as a nerve tonic and headache treatment before becoming one of the most popular fizzy drinks of all time (Lemelson-MIT, n.d.). Similarly, the term ‘cartoon’ has evolved from its initial meaning as preparation material to denote the finished drawings themselves.

With the advent of the creative renaissance, the bimodal form of visuality which includes both cartoons and texts began to grow; comics and other visual narratives are probably the most modern forms of visual arts that generally have a sequential (sometimes situational) way of conveying messages; the history of visual narratives has been through many paths in different parts of the world. Instances of such early sequential art forms can be found in Egyptian hieroglyphs, Greek friezes, Rome’s Trajan’s Column (110 AD), etc. (‘Trajan’s Column’, 2018). Certain Biblical sequential art forms were also discovered through some Christian manuscripts (Lyttle, 2015). Furthermore, manga and manhua are based on Chinese characters. The first manhua dates back to 900 years ago in the Northern Song Dynasty of China. The story portrays a bird named Manhua that catches fish with its beak sticking in the water. This influence has blended the traditional Chinese brush paints and American, Japanese and Korean comics. (Yan, 2019.) These sequential ways of art were also found to be popular in West Bengal, India as *Patua* paintings through the origin of the *Patua* tradition in the 18th century CE, though some claim that the oral form goes back all the way to the 10th or 11th century CE; the stories were often sung by performers along with the display of the sequential paintings, giving it a film-like outlook of bygone eras (Ahire, 2020). These *Patua* artists are still found in very small numbers as they are almost disappearing. Ahire posits that the artists narrate stories from epics like the *Ramayana* and the *Mahabharata*, the *Puranas*,

religious texts like the *Mangal Kavyas*, *Behula* and *Lakhinder*, the renunciation of *Chaitanya*, *Krishna Leela* and many more. According to Widewalls Magazine (2016), the very first manga in Japan appeared in the 12th and 13th centuries, in a series of drawings such as frogs and rabbits titled *Chōjū-giga*, i.e., scrolls of frolicking animals. In the 17th and 18th centuries, caricaturing and satirising of political and social issues through art started developing, which has a long tradition in Britain, dating back to the political and social shifts of The Enlightenment; prints contained several images to relate multiple scenes of a narrative. (Bury & Mellon, 2022.) Later, in the twentieth century, newspapers across cultures brought a trend of satirical drawings through cartoons and written languages (Leon, 2017): According to Cohn (2012), ‘Truly, there seems to be an intuitive link between comics and language in the minds of the creators- a belief shared by several researchers of language who, with growing frequency, are discussing properties of comics in a linguistic light.’ There is a natural relationship between comics and language. This link is underscored by numerous conversations about how comics can convey complicated ideas using both visuals and text. Therefore, comics and other visual narratives do what any other language would do—convey messages. But these visual narratives convey messages in such a way (more in an exaggerated manner) that we cannot do in our physical states: if someone is angry, probably their head would be on fire; again, if someone were in love, their eyes would literally reveal it all by turning them into heart icons! So, in this way, the visualities of comics or other visual narratives and picture stories turn out to be more prominent with a hyperbolic atmosphere, and probably that’s why these are also full of ‘superheroes’, as they bring out and embody extreme strength, speed or emotions. They bring out what exactly goes on in our minds in their crudest forms possible. They become more illustrious as they come with written languages and several visual cues, giving them a bimodal outlook and a multisensory effect, which eventually provides a sequential sense to our cognitions.

1.1.1 Visual Linguistics: A Part of Linguistic Studies

The main branches of formal Linguistics are generally categorised by language structure and language use. The structure of language mainly focuses on word structure (Morphology), sentence structure (Syntax), speech sounds and their rules (Phonetics and Phonology), and meaning in language (Semantics and Pragmatics). On the other hand, Sociolinguistics, Psycholinguistics, Cognitive Linguistics, Historical Linguistics, Stylistics, etc., fall under the category of how language is used. According to language scholar Sweet (1892), ‘Language is the expression of ideas by means of speech sounds combined into words. Words are combined into sentences, this combination answering to that of ideas into thoughts.’ Furthermore, American linguists Bloch and Trager (1942) formulated language as a system of arbitrary vocal symbols through which a social group cooperates. Therefore, from the surface level, it may seem that Visual Linguistics has nothing to do with Linguistics, but if we focus on ‘the expression of ideas’, we can say that ideas originate with the help of visualisation in our cognitions. Moreover, if we take ‘language as a system of arbitrary vocal symbols’, we can figure out that these arbitrary vocal symbols are representations of something or someone visible to us in some way. Linguists generally have viewed Linguistics from a ‘speech sound’ point of view; Visual Linguistics has always received less attention because of it. But we must investigate the matter that our lexicons are full of visual entities along with words. We all know that our lexicons are dictionaries full of words: I would rather say that our lexicons are more like picture dictionaries: a word that is present there also has a visual counterpart. Therefore, if we consider Visual Linguistics through the lens of narrative structures in visual narratives, it closely mirrors the elements found in traditional lexicons—words and images—which are innate faculties akin to language. Visual narratives help structure our thoughts by combining relevant words with sequential images. Thus, Visual Linguistics should be classified as a type of language usage under the field of Linguistics. Furthermore, given its grammar and structure, Visual Linguistics should also fall under the broader category of language structure.

1.1.2 Visual Linguistics and The Theory of Innateness

The frontal lobe of the human brain is the largest; it is very vital to our consciousness, and its functions are uniquely human, as it is responsible for language (Villines, 2023). It is much more advanced in cognition than that of other animals. Several scholars, right from Socrates, Plato and Aristotle, tried their best to deal with this subject. Though Plato (1943) did not mention anything about human cognition, yet, it seems that his ‘mimesis’ tries to explain human cognition through the theory of imitation, and later it was explained by Aristotle with more details. It explains that every art is mimetic in nature, therefore being an imitation of life. It was theorised that the ultimate reality is an ‘idea’ that takes birth in human cognition. Art imitates these ideas, thus becoming the imitation of reality. Plato gives an example of a carpenter and his idea of a chair that came to his mind. The carpenter gave physical shape to his idea; an artist then imitated this physical shape in his picture, thereby giving birth to the third imitation. Thus, the chair gets removed from ‘reality’ twice. This theory can be linked with the mental representations or the formation of lexical rules. ‘Perhaps the most frequently cited aspect of Saussurean doctrine is that a word is an arbitrary association of a sign with a meaning’ (Jackendoff, 2003, p. 51). One of the most widely quoted aspects of Saussure’s (1916) linguistic concept is his claim that a word is an arbitrary combination of a symbol with a meaning. He contended that there is no inherent or natural relationship between the sound of a word (the signifier) and its meaning (the signified). Instead, this relationship is based on social conventions. For instance, the word ‘tree’ refers to a specific type of plant, not because the sound ‘tree’ resembles a tree, but because English speakers have collectively agreed on this association. This is similarly found in other cultures, too. This notion emphasises the arbitrary character of linguistic signs and the importance of social constructs in language; here comes the word-meaning relationship. People use words to communicate and convey certain meanings. When people see a pot, they would call it a ‘pot’ to communicate with other speakers and listeners. Similarly, when artists see a pot, they imitate the pot and draw it.

This is how an artist communicates through their artwork with the world. When we see an artwork, our brain automatically comprehends the content of the artwork. By recognising the several shapes and spatial relations of different objects, we can cognitively connect to the artwork. These shapes and spatial relations are nothing but representations of the natural objects present around us. Therefore, visual art is nothing but our perceptions of how we see the world and how we think: through these artworks, the artist communicates with the audience, their fellow viewers. This type of communication is termed visual communication. Written language evolved much after the cavemen had already started communicating through cave paintings, and sequential images have been a natural means of conveying narratives (Cohn, 2013). Now, in the modern world, picture stories or visual narratives are quite in vogue—the quest here creates a thin mental line between the word created and the mental representation of it.

Everyone tells a story inside their mind, and authors write them out with written language. So, what do the artists do? They tell their stories, too, through their medium of communication. A picture can tell a story exactly like a book. If we consider visual narratives as references, these stories become even more prominent and communicable. Several illustrators or authors of visual narratives have tried to prove and mention the properties of comics in different forms as a language. We start acquiring language after our birth (Chomsky, 1957); comics are also acquired gradually as we learn to decode visual narratives, but the question that arises is, do we acquire the grammar and structure of these picture stories the same way we acquire the knowledge of our mother tongues? We probably visualise it and then fit it into our cognitions according to our comprehension, depending on what culture we belong to. Visual narratives are not works of intuition; the properties of comics work in such a way that they act like the properties of language, as McCloud (1993) has described the sequential panels of comics as its grammar. According to Cohn (2013), ‘Ultimately, the definition of comics includes a network of ideas composed of their subject

matter, format, readership, history, industry, the specific visual languages they use, and other cultural characteristics.’ (p. 2). Though we do not acquire comics the same way we acquire our mother tongues or learn our second languages, it has all the properties of Linguistics, be it from the syntactical, semantical, or pragmatical point of view. The central enigma of this research lies in uncovering how the two selected cultures perceive visual narratives differently, with a particular focus on the cross-linguistic dimensions revealed through both the differences and similarities in their visual linguistic structures. Comics and other visual narratives are forms of art that have the power to create sequential stories, and amazingly, they work like movie scenes or films in the human brain. Visuality opens several grounds as it is directly connected to the brain. A University of Utah study on brain lateralisation examined the brains of more than 1,000 people and found no evidence of people having any dominant side (Nielsen, et al., 2013). Even then, each side of the brain does more work for certain functions: it is argued that our left side of the brain is all about language and it consists of orders, logic, symbols, and objectives, whereas the right side of the brain is all about creativity, consisting of art, innovations, emotions, and colours (Cherry, 2024). It seems that picture stories can bring both these sides of the human brain together, making logic and creativity meet. If the whole brain is engaged through a single medium like visual narratives, does that suggest that Visual Linguistics provides a powerful means to perceive and understand the sequential flow of language—something inherent to us? Or is it the other way around—do we understand things better when visual elements are involved? This is probably why visual narratives differ across cultures—we tend to interpret them through the lens of our own cultural perspectives.

1.1.3 Cross-Cultural Perceptions in Visual Linguistics

According to Cohn (2012), ‘The very notion of cultural drawing styles belies a purely perception-based model of drawing because styles are built from conventional patterns shared by people of a culture.’ In this context, this dissertation focuses on investigating how human

cognition, influenced by social contexts, operates similarly and differently across cultures and over time when perceiving different types of picture stories, even those that are culture-specific. Cohn (2013) specifies that ‘Drawings and sequential images are an integral part of human expression dating back at least as far as cave paintings, and in contemporary society appear most prominently in comics.’ The visual forms of communication have grown substantially over time, and in the present society, visual narratives are the most notable ones. Such narratives use images and texts to create interesting and dynamic stories that make difficult events and concepts understandable to a wide range of audiences. This indicates that it is an effective tool for cultural expression, education, and entertainment, demonstrating the lasting importance of images in human communication. By combining the vocabularies present in our brains, the artist brings out the creative ability and forms meaningful units in their comics. Therefore, does it differ cross-culturally and across time?

Language is diverse; so is Visual Language. As mentioned earlier, human beings are born in different cultures with plenty of diversity, yet again, we should have some universalities among us, especially cognitively. This is because human beings have an innate ability for visuality. Chomsky’s (1977) hypothesis postulates that children are innately born with a sense of language and grammar. It seems that we live in one reality, and another reality resides in our brain. It should also be a matter of fact that whatever we have in our inner mind or in our cognitions is an influential subject that we get from our surrounding society or culture. So, even if a person is blind, they would be able to visualise or conceptualise a figure with the help of braille through the Visual Linguistics that is present in the inner reality of the brain: does this mean that human beings are not only born innately with the sense of language and grammar but also with the social surroundings that they are culturally born with?

Every person has their own style of speaking and writing when it comes to language: the unique speech patterns or language habits characteristic of an individual are linguistically referred to as their ‘idiolect’. Art through comics also takes the essence of this idiolect and

makes every illustrator or artist unique. Every comic artist has their way of communication through this unique blend of images and texts (sometimes only images), and this goes deep into human cognition, where it would mainly focus on the philosophical choices of Visual Linguistics from a macro point of view. In this regard, it could be said that Sociolinguistics and Cognitive Linguistics are inter-related with each other when it comes to Visual Linguistics and the Chomskian Theory of Innateness is also a part of Visual Linguistics; cultural influences, education and other socioeconomic factors help us to pass perceptual information through cognition and how we perceive across time and culture is what this research would mainly focus on, with selected references.

The research aims to explore the criteria along which visual culture differs from one another, with special reference to two different forms of visual narratives (Bengali comics and Japanese manga) and their structures as a form of visual representation. Cross-cultural differences happen due to the variations in the visualisation of things and events, not only from the viewpoint of what they are in themselves, but also from the viewers' interest in knowing how they are structured and visually communicative. The basic question which triggers the mind as a researcher here is—how do things and events possess any intrinsic true essence, while there remains an important space for cross-cultural variations since ancient days and how does it affect human cognitions (similarly or differently)? The two visual narratives, *Hada Bhoda* and *Doraemon*, differ in various ways, reflecting how artists from the same cultural background can approach drawing differently. At the same time, viewers across cultures and time periods may perceive and decode the same artwork in similar ways, influenced by shared cognitive and cultural frameworks.

Because of cultural differences, comprehension as well as production of visual narratives often seeks to make a series of aesthetic and philosophical choices. A careful portrayal of the visual representation is proven to be of immense significance because it gives important clues about the way it is represented. It is construed by the choice of values for the parameters

under the influence of universal principles of human perception and cognition; and if there are some universal principles, then there should be an innate ability to understand the visual narratives produced across the globe. Therefore, if there is some kind of basic or intrinsic essence like our innateness, then how do these visual narratives differ from one culture to another?

Cross-cultural communication is the study of how people communicate in different cultures; how they differ and how they are similar. With the spread of global businesses, the need for cross-cultural communication has been increasing (Matthews & Thakkar, 2012). There is no doubt that the digital age has changed how we generally view communication. Transforming ideas using symbols and signs is visual communication. Using visual data is not an uncommon factor among presenters in various fields. These visual data are pretty much comprehensible and acceptable among people cross-culturally. Therefore, reading visual narratives gives us a unique experience of visual communication. Lloyd (2015) states that visual language is 'the language of images.' Visual Linguistics is non-verbal communication, and it allows the illustrators of visual narratives to communicate with the readers. The primary focus of Visual Linguistics in comics and graphic stories is mainly on the images, and these images are much more powerful than the texts. Readers of comics may respond differently according to their age or the culture they belong to. As stated by Veerasarn (2018), 'the publishers of the special issues stated, the readers might still understand the content, but the emotional response might be different from the current generation, given those who were raised under the King's reign are attached to Rama IX and feel more strongly than those who do not have a firsthand understanding of the King.' It is evident from this that the younger generation who were not raised under King's rule did not have the same emotional response as those who grew up during his reign. Essentially, the emotional impact of the content depends on whether the readers have personal memories or experiences tied to the King, creating a generational divide in how the same material is perceived emotionally.

The way people emotionally respond to content is often shaped by their personal experiences, which is why the same material might affect different generations in distinct ways. This becomes clearer when we think about cartoon images, which are essentially symbolic representations of real-life people, animals, and objects. The emotional impact of these symbols can vary greatly depending on how familiar or connected a reader is to the real-world references behind them. The question that arises here is—could these illustrations be interpreted similarly or differently across cultures or across time, or even differently within a specific culture, depending on human cognitions? For instance, the use of certain colours, like red, in images may ignite a political viewpoint in readers irrespective of their cultural backgrounds. Likewise, a facial expression depicted by a cartoon character may be unfamiliar to a reader from a culture different from that of the artist or illustrator. Translating graphic narratives may lead to onomatopoeic (idiophones) issues across cultures, but sometimes it also enhances creativity rather than killing it. Every human being perceives an artwork in their own unique way through life experiences and personal choices. Even though our lexicons are full of images and words, our perceptions may vary. The central theme that would be seen throughout this research is how meaning is communicated through the Bengali comics, *Hada Bhoda* and the Japanese manga, *Doraemon*, cross-culturally, and the structural, grammatical, and social contexts are looked into through these picture stories. The temporal aspects are also explored with the help of Bengali *Patua* and Japanese *Emaki* scroll paintings, respectively.

McCloud's (1993) reflection on the 'strangeness' of comics captures the complex interplay of visuals, cultures, and linguistic features that make the medium of visual narratives unique. His 'tick' in this research is analysed through the differences and similarities in how the selected visual narratives from both cultures of Bengal and Japan highlight the linguistic elements, highlighting the fact that visual narratives convey culturally embedded messages and not just universal truths. For example, certain visual elements may substitute for complex

events or emotions, but are placed within a textual and visual sequence according to cultural conventions. The exploration of diachronic and synchronic variations through the variations of Bengali and Japanese scroll paintings further underlines that these ‘ticks’ are often historically embedded, contributing to each culture’s uniqueness of storytelling. By emphasising the fact that culture organises and interprets these visual features, the dissertation aligns with McCloud’s notion of comics as a ‘strange’ and complex medium. It is not merely a difference in artistic arrangements but also in the Linguistics of these cultural visual narratives that bring out the distinctive ‘ticks’ of these cultures:

‘I’ve been trying to figure out what makes comics “tick” for years and I’m still amazed by the strangeness of it all.’ (McCloud, 1993, p. 117).

1.2 Statement of the Problem

In recent years, visual narratives, especially comics and scroll paintings, have gained scholarly attention as rich semiotic artefacts. Nevertheless, much of the existing literature has focused on either formalist narrative structures (McCloud, 1993) or the assumed universality of visual language (Cohn, 2013). These studies tend to treat visual storytelling as a system that transcends linguistic and cultural boundaries. While these frameworks have proven useful in analysing visual grammar and cognitive processing, they often overlook how deeply cultural contexts shape the way stories are visually told and interpreted. As scholars like Forceville (2009) and El Refaie (2009) argue, the assumption that images speak a universal language can obscure the role of localised conventions, traditions, and interpretive frameworks.

Under this light, this dissertation delves into the implications to explore how cross-cultural visual narratives represent different cultures and bring out a communication system within the medium. The intersection of Linguistics and picture stories presents a complex problem, giving rise to curiosity about how structural, functional, diachronic, and synchronic aspects of

Linguistics fit into the analysis of visual narratives, as well as what role speech acts through linguistic pragmatics play in shaping the meaning of picture stories across cultures.

Although recent works have started to explore cultural variations in visual narratives, particularly in the context of Japanese versus American comics (Cohn et al., 2012), such efforts remain partial and are largely limited to popular manga and American superhero comics. Few studies have undertaken comparative analyses of traditional visual storytelling forms like *Patua* scrolls or *Emaki-mono* alongside contemporary comics. Moreover, the role of linguistic theories in explaining how meaning is constructed within visual narratives across cultures remains significantly under-theorised. The communicative strategies embedded in panel sequencing, temporal aspects, facial expressions, ideophones, and drawing styles, though recognised, are rarely linked back to structural, functional, diachronic, or pragmatic aspects of Linguistics.

Therefore, this research identifies a critical omission in the field: the inadequate application of linguistic frameworks to the study of cross-cultural visual narratives. For example, while Chomsky's (1957) structural approach has informed sentence-level syntactic analysis, its potential to decode the underlying structure and grammar of visual sequencing in culturally specific ways has rarely been explored. Similarly, Halliday's (1961) functional theory, particularly its emphasis on how language serves social purposes, can provide insights into how visual narratives reflect and perform cultural roles, yet remains largely unutilised. Saussure's (1916) synchronic and diachronic approaches offer valuable tools for exploring temporal shifts in visual language, especially in traditional forms like scroll paintings. Lastly, Austin's (1959) speech act theory, though extensively applied in verbal communication, has yet to be fully explored in the context of multimodal storytelling, particularly in how facial expressions, gestures, and other visual symbols function as communicative acts in picture stories. The theories of visual narratives present a complex problem, as the thematic coding

of visual narratives through the linguistic aspects across cultures remains understudied. The different linguistic aspects, as mentioned, may impose a latent thematic structure on visual storytelling, influencing their meaning and interpretations across cultures. Thus, this study seeks to fill this theoretical gap.

The prevalent assumption of a universal essence in visual storytelling gives a murky interpretation among cross-cultural contexts, linguistic conventions, and different aspects of picture stories. This limits the understanding of how culturally specific visual entities construct meaning, brushing aside the nuanced manners in which cultural frameworks influence panel sequencing, drawing styles, and other social interactions. Moreover, the failure to account for the cultural and temporal contexts of traditional visual storytelling in the form of scroll paintings may restrict the understanding of the dynamic aspects of universal cultural exchange of its social impacts. Therefore, this research addresses this gap by analysing the intricate relationships among cross-linguistic variations through linguistic studies and narrative structures.

Understanding cross-cultural variations is crucial for bringing out the nuances of communication and meaning in diverse cultural settings and contexts, making it a deeper level of human cognition for complex relationships between language, culture, and visual storytelling.

The study seeks to address the pressing issue of whether and how the cultural parameters play a significant role in shaping visual narratives, resulting in varied, heterogeneous visual cultures across different societies. This research particularly aims to explore the cultural variations with the help of different linguistic aspects by comparing the two selected picture stories, *Hada Bhoda* and *Doraemon*, as well as the traditional forms of visual storytelling, such as *Patua* and *Emaki* scroll paintings from the two distinct cultures of West Bengal and Japan, respectively, to identify cultural factors that may influence the styles of visual narratives.

To reveal how deeply culture informs not just what is told but how it is visually expressed, this study was driven by the need to critically analyse whether visual narratives carry hidden cultural codes that influence perception, emotion, and meaning. By doing so, it challenges the notion of visual universality and seeks to contribute to broader discussions in linguistics, semiotics, and intercultural communication through the lens of visual storytelling.

Therefore, it seems that the pervasive presence of cross-cultural differences in visual narratives poses a fundamental problem—if things and events have any intrinsic true essence, there would have been no space for cross-cultural variations; there would be no room for cross-cultural variations if things and events had an intrinsic, genuine nature. The prevalence of cross-cultural variations may suggest that the picture stories are not just objective reflections of reality, but also subjective interpretations that can be filtered through cultural perspectives. This raises the curiosity of the research as well as certain critical objectives, which will be revealed in Section 1.3. As such, the nature of visual representation, the influence of culture, and the construction of meaning can be effectively understood through the analysis of linguistic elements.

The complexity and richness of visual narratives across cultures underline a deeper, unresolved problem at the heart of Visual Linguistics: the nature of meaning itself. If images are indeed universal, capable of conveying ideas independent of linguistic and cultural boundaries, then the considerable differences observed in the storytelling techniques, styles, symbolisms, and interpretive cues across cultures pose a significant contradiction. The very existence of these cultural variations calls into question the universality of visual language and, by extension, suggests that meaning is not inherent in images themselves but is constructed through the lens of culture, experience, and context.

This contradiction brings forth a critical tension between two poles of understanding: one that assumes images, like language, follow a set of universally comprehensible structures, and

another that argues meaning is deeply contingent upon cultural conventions, historical experiences, and socio-political realities. If visual narratives were purely objective, then readers from different cultural backgrounds would interpret them in largely similar ways. However, the diversity of visual storytelling practices, from the linear, often comedic realism of Bengali comics like *Hada Bhoda* to the technologically imaginative and emotionally expressive world of Japanese manga such as *Doraemon*, shows that visual literacy is not uniform. Rather, it is learned, practiced, and embedded within the cultural contexts that shape our interpretive habits.

Furthermore, these differences are not trivial. They encompass not only artistic techniques, such as panel layout, facial expression conventions, and the use of ideophones, but also more profound sociocultural values, such as representations of family, education, social harmony, etc. This becomes particularly problematic when scholars, educators, or even global audiences approach visual narratives with the assumption that images ‘speak for themselves’. This research challenges such assumptions by arguing that images do not simply mirror reality; they construct it. This construction is filtered through the visual grammar specific to a culture, just as spoken or written grammar reflects the unique syntax and semantics of a language community.

This raises the central concern of the present research: How can we understand visual narratives in a way that accounts for both their structural components and their cultural embeddedness? How can we approach images not merely as aesthetic objects but as forms of discourse that communicate, persuade, and construct meaning in culturally specific ways? If meaning is not fixed in the image but negotiated through a shared set of signs and codes, then visual narratives must be treated not just as stories, but as acts of communication shaped by the social, historical, and cultural contexts of their production and reception.

The pervasive presence of cross-cultural differences in visual storytelling does not represent a

barrier to understanding but rather a site of rich inquiry. These differences force us to confront the limits of our own interpretive frameworks and challenge us to develop new, more inclusive methods of visual analysis. They also remind us that what is seen, and how it is seen, is never neutral. Every visual narrative encodes a perspective, and every act of interpretation is shaped by the viewer's own cultural lens.

In the absence of such a multifaceted approach, the study of visual narratives risks remaining fragmented, either leaning heavily on literary narrative theory or on cognitive and formalist accounts that do not adequately account for cultural variation. The core contradiction this research addresses is the tension between the assumed universality of images and the cultural specificity of visual meaning-making. If visual language were truly universal, we would expect a high degree of interpretive uniformity across cultures. Yet, as evidenced by the contrast between the comedic realism of Bengali comics like *Hada Bhoda* and the emotionally charged, technologically imaginative world of Japanese manga like *Doraemon*, this is clearly not the case. Moreover, the use of traditional storytelling forms, *Patua* and *Emaki* scrolls, illustrates not only diachronic shifts in visual narrative conventions but also culturally specific representational choices that remain largely unstudied.

Thus, the problem this dissertation addresses is not merely how images communicate, but how they come to mean differently across cultures. It is a problem that invites interdisciplinary exploration, drawing on linguistics, semiotics, cultural studies, and visual anthropology. It asks us to rethink what we consider to be the 'language of images', and to recognise that this language is not static or universal, but perhaps dynamic and negotiated.

The variability in visual narratives across cultures challenges the assumption of visual objectivity and calls for a more nuanced, culturally sensitive approach to the analysis of picture stories. This research takes up that challenge, arguing that understanding the structure, grammar, and function, along with temporal aspects and different communicative acts of

visual narratives in context, is essential not only for appreciating their aesthetic value but also for grasping their role in cultural communication, identity formation, and meaning-making.

1.3 Objectives and Research Questions

Building on this statement of the problem, this dissertation addresses that images do not communicate in a fixed or universal way; rather, they acquire meaning through cultural and temporal contexts. Therefore, this study aims to explore how such meanings are constructed and vary across societies and time. Recognising that visual narratives operate within culturally shaped ‘languages’ that are negotiable and dynamic, this research sets out to analyse how picture stories reflect and reinforce distinct social norms, values, and perceptions. Accordingly, the general objective of this dissertation is to examine cross-cultural perceptions through the selected visual narratives.

This dissertation examines the complex relationships of cross-cultural themes in visual narratives through *Hada Bhoda*, *Doraemon*, *Patua* and *Emaki* scroll paintings from a linguistic point of view. This research is noteworthy because it exposes how different cultures and social norms are represented linguistically and how they interact within the medium of visual narratives, providing insights into global cultural exchanges and the impacts of visual narratives on societal perceptions.

Under this, my general objective is as follows:

To study cross-cultural perceptions through selective visual narratives and explore if and how the temporal variation in the communication of visual culture varies over time.

To specify the general objective, the research came up with three discrete objectives: these objectives are based on the two cultures that are selected for the study: West Bengal and Japan. These two cultures are selected since I belong to the Bengali community; therefore, it would be easier for me to compare. Secondly, Japan is selected because it has one of the

richest cultural backgrounds regarding art and visual narratives, similar to West Bengal. The Bengali comics, *Hada Bhoda*, and the Japanese manga, *Doraemon*, were chosen as my stimuli for this study because of their popularity, rich visual and narrative contents, various interesting representations, and cross-cultural appeal with social influences. It should also be noted that even though they are similar in certain ways, there are a wide range of differences between these two visual narratives, which will be unfolded in the chapters of the findings. The characteristics of these two visual narratives make them effective tools for exploring and interpreting cross-cultural issues, which effectively aligns with the goals of this dissertation. Some *Patua* scroll paintings are also used in this research, which I collected myself as data, along with a few *Emaki* scroll paintings found on the internet for the study as the stimuli to see how Bengali comics and Japanese manga came into being, respectively, since ancient days.

Therefore, the three specific objectives are as follows:

1.3.1 Objective 1

To cross-culturally analyse the visual narratives of *Hada Bhoda* and *Doraemon* through the grammar, structure, and social contexts of visual culture used in these visual narratives.

Under this, my key research questions are as follows:

- (i) How are the panels ordered? Do they have any cultural importance?
- (ii) What are the differences in illustrations? How do the narrative structures vary between these two cultures?
- (iii) How do the social contexts of the visual narratives differ between these two cultures, depending on the social surroundings of family dynamics, friendly exchanges, school conversations, sportsmanship and culinary culture?

1.3.2 Objective 2

To find out how the structures of *Patua* scroll paintings and *Emaki* scroll paintings differ from the visual narrative grammar of Bengali comics and Japanese manga, respectively.

Under this, my key research questions are as follows:

- (i) How is the iconicity of the achronological panels ordered?
- (ii) How are the natural gutters (in scroll paintings) semantically different from what they are today?
- (iii) Which events are illustrated and which are not, and how have they changed in their social contexts over time?

1.3.3 Objective 3

To explore how visual culture is communicated in the visual narratives of *Hada Bhoda* and *Doraemon* through different communicative acts, facial expressions, ideophones or onomatopoeic words and drawing styles.

Under this, my key research questions are as follows:

- (i) How is the visual culture conveyed through the different communicative acts and facial expressions of the characters in the two visual narratives?
- (ii) How are the ideophones conveyed in these visual narratives? How do they differ from one another?
- (iii) What are some prominent and different drawing styles across these two cultures?

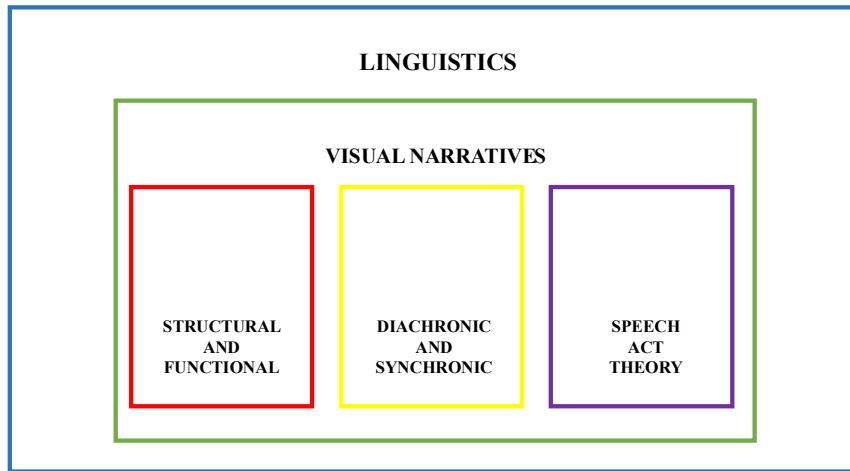
Since cultural nuances are there, comprehension as well as the production of visual narratives often seek to make a series of aesthetic and philosophical choices. A careful rendition of the visual representation is proven to be of immense significance because it gives important clues about the way ‘represented’ is construed by the choice of values for the parameters under the influence of some universal principles of human cognition.

This study traces its origin back to the prehistoric era of human civilisation. The fundamental goal of this research is to investigate and analyse the cross-cultural perceptions of visual narratives through the objectives with the help of *Hada Bhoda* and *Doraemon*, mainly along with the *Patua* and *Emaki* scroll paintings, respectively, for the temporal variations. A brief overview of the background, along with the significance and cross-cultural relevance of this study, has been presented in this chapter, exploring the distinct traits of visual narratives and how they are perceived by different cultures. Additionally, this chapter has also outlined the initial skeletal theoretical framework, guided by the objectives of the study, to depict how different cross-cultural perceptions are analysed with the help of cross-linguistic features.

1.3.4 Theoretical Foundation Guiding the Objectives for Cross-Linguistic Analysis

A theoretical framework based on foundational theories from Linguistics supports these objectives and provides a basis for cross-linguistic analysis of the selected Bengali and Japanese visual narratives. This section of the chapter gives an outline of the initial blueprint of the framework, which will be discussed in detail in the next chapter. This framework includes Chomsky's (1957) Structural Theory, Halliday's (1961) Functional Theory, Saussure's (1916) Diachronic and Synchronic Analysis, and Austin's (1959) Speech Act Theory. Each of these theories aligns closely with one of the specific objectives guiding the analysis of narrative grammar, structure, and functions, temporal variations, and the interplay between textual and visual elements through communicative acts in the cross-cultural contexts of West Bengal and Japan.

Figure 1. Initial Framework



Source: Developed by the author (Chattopadhyay, 2023)

Building upon this framework, the thematic coding of the study is employed as an analytical tool to systematically categorise and bring out the visual linguistic elements within *Hada Bhoda*, *Doraemon*, and the selected *Patua* and *Emaki* scroll paintings. Therefore, this enables a cross-linguistic analysis of the narrative structures with the help of the different linguistic aspects discussed through the framework. Each theoretical perspective directly informs the coding themes that provide a linguistic structure to analyse how elements in visual narratives, such as panel ordering, social interactions, temporal variations, and communicative acts through ideophones and other artistic styles, vary between Bengali and Japanese narratives. This alignment ensures that the thematic coding process is grounded in theoretical precision and tailored to unveil these visual narratives' cultural nuances and thematic contrasts. This approach, as a result, extracts the key cross-linguistic insights by categorising them accordingly with the help of the linguistic aspects, allowing for a comprehensive analysis that supports the objectives of this study. A content analysis methodology was applied to explore these elements systematically to achieve this, which is discussed in detail in Chapter 3. This approach, as a result, extracts the key cross-linguistic insights of the study.

1.4 Significance and the Study

Under this light, this study makes significant scholarly and practical contributions across several dimensions, particularly to the emerging field of Visual Linguistics, while also offering insights applicable to pedagogy, cultural studies, psychology, and media or communication studies. Its value lies not only in bridging cultural perspectives but also in establishing an innovative interdisciplinary methodology that synthesises linguistic theory with visual narrative analysis.

(i) Disciplinary Contribution and New Knowledge in Visual Linguistics

The primary academic contribution of this research lies in extending the domain of Visual Linguistics by demonstrating how structural, functional, diachronic, and pragmatic linguistic theories can be systematically applied to analyse cross-cultural visual narratives. While existing scholarship has acknowledged the semiotic richness of comics and visual storytelling (McCloud, 1993; Cohn, 2013), few have grounded their analysis in mainstream linguistic theory. This dissertation uniquely integrates the insights of Chomsky's (1957) syntactic theory, Halliday's (1961) systemic functional grammar, Saussure's (1916) diachronic/synchronic perspective, and Austin's (1959) speech act theory to interrogate visual texts across cultures. By doing so, it generates new theoretical and analytical pathways for the field of Visual Linguistics, which remains underexplored in this context.

The study presents a cross-linguistic analysis of the visual grammars and other communicative strategies in *Hada Bhoda*, *Doraemon*, and temporal shifts in traditional scroll paintings (*Patua* and *Emaki*), offering empirical insight into how language-like features, such as sequencing, diachrony-synchrony, ideophones, facial expressions, and communicative acts, operate differently across cultures. This addresses a recognised lacuna in existing literature, which tends to focus on either American or Japanese visual systems in isolation (Cohn et al., 2012). The research, thus, contributes new knowledge by establishing cross-

cultural, linguistically grounded frameworks for interpreting visual narratives.

(ii) Pedagogical and Policy Implications

This study has strong implications for curriculum design, educational policy, and multilingual literacy development. The findings advocate for the inclusion of culturally embedded visual narratives in language and literacy classrooms. Visual texts accompanied by culturally specific narrative structures and linguistic elements can enhance intercultural competence, develop critical thinking, and support metacognitive awareness in learners. The study positions comics and traditional picture stories as pedagogical tools capable of fostering deeper engagement and contextualised learning, especially in multicultural and multilingual settings (Chattopadhyay et al., 2024).

In policy terms, this dissertation argues for the integration of visual literacy and cross-cultural narratives into national curricula, especially in South Asia and East Asia, where traditional art forms and comics are still sidelined in formal education. The emphasis on decoding images as cultural texts can enrich language teaching policies and promote inclusive, culturally responsive education.

(iii) Cross-Disciplinary Application: Psychology, Media, and Medicine

The implications of this research also extend into psychology, psychiatry, media studies, and health communication. Creative arts are increasingly being used in therapeutic contexts, especially in psycho-medical fields such as trauma counselling, autism education, and mental health literacy (Frasco et al., 2025). The culturally informed reading strategies developed in this study can help psychologists and clinicians design more effective, context-sensitive visual interventions. Therefore, visual narratives can also play a crucial role here. Furthermore, in media and communication, understanding how different cultures construct and interpret visual content can also inform cross-cultural media production, advertisement design, and audience engagement strategies.

(iv) Preservation and Recontextualisation of Cultural Heritage

This research contributes to the documentation, preservation, and reinterpretation of traditional art forms such as *Patua* and *Emaki* scroll paintings, which have received limited analytical attention in academic literature. By comparing them with contemporary visual narratives, this study not only maps their evolution but also contextualises their socio-political and narrative relevance in modern times. It brings marginalised or regionally rooted storytelling practices into academic discourse, emphasising their cultural value and narrative intelligence.

(v) Promoting Intercultural Understanding

Finally, this research plays a vital role in fostering intercultural dialogue. By exploring how picture stories from West Bengal and Japan express social norms, identities, and emotional landscapes, it deepens our comprehension of how cultures shape the way we see, feel, and narrate. The findings of this study can thus support educators, cultural practitioners, and policy-makers in designing intercultural learning frameworks that are grounded in both visual and linguistic literacy.

Therefore, by offering a linguistically informed, cross-cultural investigation into visual narratives, this study fills significant theoretical, analytical, and pedagogical gaps in Visual Linguistics. It not only advances academic understanding but also provides practical frameworks for educational, clinical, and cultural applications. Its interdisciplinary approach makes a compelling case for recognising visual narratives as serious cultural texts, worthy of scholarly engagement and practical application in a globalised world. Hence, this study makes a robust and original contribution to the linguistic field and beyond by paving the way for future interdisciplinary research at the intersection of language, culture, and visual storytelling.

1.5 Key Operational Terms and Narrative Frameworks in Visual Narratives

Comics create a film-like narration in the human brain when read. It presents a sequential way of understanding the whole narrative with pictures and texts, and sometimes with only pictures. Postema (2013) asserts that her ‘study interrogates comics as comics, in all many genres and with a focus on the formal and material specificities of the form.’ She questions in her study how comics create their narratives. Comics, as a medium, has evolved in both form and function, shifting from simple entertainment to a powerful tool for social commentary. ‘Like the divide between “comics” and the “comics medium,” “art” can be understood as a social term applied interpretively to varying actions and objects.’ (Cohn, 2005). Therefore, the notion of art is fluid and is influenced by societal values on subjective perspectives, depending on what is considered meaningful and aesthetically significant. Western comics, graphic novels, and other comics that follow the Western pattern are read from top to bottom and left to right. The sequential panels create a story in the readers’ brains. On the other hand, manhwa, manhua, and manga originate in eastern Asia. Manga refers to Japanese comics while manhwa and manhua refer to Korean and Chinese comics, respectively (Vergara, 2021.) Unlike Western comics, manga and manhua are read from top to bottom but right to left. However, manhwa is read from top to bottom and left to right just like Western comics. Nevertheless, the horizontal way of looking at comics changed with the advent of smartphones, and manhwa is often read vertically with the scrolling system; this format is referred to as webtoons. Mangas are generally black and white and have multiple panels. Manhuas are mostly colourful and have a single-issue format and on the other hand, manhwas are the recent forms that have a balanced blend of Western comics and East-Asian styles (Peralta, 2020.)

Everything that is worthy of reading a book may also be worthy of reading a comic book. Visual narratives could be understood by the grammar that is embedded in them, similar to the studies in Linguistics. McCloud (1993) describes closure as, ‘Observing the parts, but perceiving the whole.’ This is very important for picture stories when it comes to their

narrative structures because the actions that are read by the readers do not come in real-time. The author can draw a static picture, and the reader can only read a static picture at a time. Just as traditional books can convey complex and meaningful content worthy of serious reading, visual narratives, too, can possess their own linguistic elements that guide interpretation, much like linguistic grammar in spoken or written language. Just like Linguistics deals with language components to convey meaning, visual narratives unveil how the visual elements interact to tell a coherent story with all the visual linguistic features present in them. By ‘observing the parts, but perceiving the whole,’ readers can connect individual panels or images to create an easily flowing and comprehensible narrative, filling in the gaps between them based on context and prior knowledge. In comics, ‘closure’ allows readers to infer actions, emotions, and events that are not explicitly shown, making the readers active participants in constructing the stories. This is not unlike the readers of traditional texts, who interpret sentences and paragraphs, building meaning from individual components to comprehend the complete message. Therefore, visual narratives also bring out distinct linguistic features through the grammar that are embedded within them.

So, to convey the passage of time and movement, they come with certain specific terms significant to understanding the film-like sequences present in visual narratives; these terms build up the very narrative structure and layout of the Visual Linguistics of different forms of picture stories:

Panel: These are geometric-shaped structures that contain scenes from the visual narratives. It is an organisational framework of picture stories that helps us to understand the sequence of the whole narrative. One panel could be correlated to one paragraph of a book. When we find a sequence of these panels, we start to understand the visual narrative, one by one.

Gutter: It is the space between two panels; sometimes referred to as a ‘breaks’ or a white space. Without this gutter or space, picture stories would appear vague because it is here that

the reader makes meaning of the transition that takes place from one panel to the other. Gutters could be analogous to the falling of curtains after each scene in a play. They can represent a change of place or the passage of time, or any kind of movement, etc.

Narrative/ Caption/ Textbox: These boxes are generally found at the corners of panels, and they provide some narrations that accompany the dialogues in a graphic story. They basically focus on the space, time, and place of the scene, exactly like the setting of a play.

Dialogue Balloon/ Speech Bubble: The characters in comics and other graphic narratives use dialogue to talk to one another. These are like ‘bubbles’ that look like clouds floating above the characters’ heads. Sometimes these bubbles are followed by smaller bubbles, which indicate that the character is thinking rather than saying something aloud.

Sound Effect: The use of onomatopoeic sound effects (idiophones) is an incredibly unique characteristic feature of visual narratives. They serve as sounds that are typically written as stylised words across a panel. For example, the sound of an explosion like ‘boom’ could be written in such a way that it echoes in the ears of the readers. This is a remarkably interesting factor of visual narratives, and this can vary from one culture to another, as onomatopoeic sounds differ cross-culturally. For example, the onomatopoeic sound of a ‘sneeze’ varies across cultures.

With the help of these terms, picture stories establish themselves as a powerful medium of communication, thereby falling under the category of Visual Linguistics through syntactical, semantical, and pragmatical understanding, by decoding and making sense out of the fragments and creating a film-sequence-like essence in our brain. We all know that we have five sensory organs: eyes (vision), ears (hearing), skin (touch), nose (smell) and tongue (taste). Vision plays a dual role: on one hand, it refers to what we see with our eyes, where the retinas capture information and send it to our brain; on the other hand, it involves what we

imagine or visualise in our minds, using our 'inner eyes.' In both cases, vision shapes how we perceive the world. In a pioneering paper, American researchers Dement and Kleitman (1957) examined the relation between the movement of our eyes and the dream contents while we sleep. The study found that participants who were woken during REM sleep had a higher rate of dream recall, while those awakened at other stages showed a lower rate of recall. Moreover, they observed a connection between specific eye movements and dreaming, which confirmed this link in both healthy and schizophrenic individuals, respectively. This gives us the idea that we experience stories in our minds through visualisation, and it relates to how our brain actively processes information during narrative consumption. Therefore, we can relate visual narratives to these sequentially visualised dream stories, strengthening the point about the 'innate ability' of sequential visualisation. We can connect visual narratives to this sequential visualisation of stories, which strengthens the very point that the 'innate ability' for sequential visualisation is a fundamental element of human cognition. Visual narratives take advantage of our natural ability to comprehend information in a linear, step-by-step fashion by breaking complicated ideas into a series of images that are inserted in panels. This improves the efficiency of storytelling in visual narratives and other media by tapping into our natural ability to process and remember information in sequences.

McCloud (1993), in his 'Understanding Comics', states that comics are a part of our historical tradition. It gives an in-depth look at the overall formal aspect of comics. He further discusses that the fundamental vocabulary and theoretical ideas constitute comics as a form of art and a medium to convey ideas and thoughts. He also talks about comics as being an ancient way of conveying messages, although it has a very modern outlook in the present time. Furthermore, Will Eisner is among the first of many to regard comics as a serious art form. He discusses the foundations of comics and how to build a visual narrative that can make its mark in the field of Linguistics—no wonder the most prestigious awards in visual storytelling honour his name!

1.6 Outline of Dissertation

Chapter 1 introduced the field of Visual Linguistics and examined the different cross-cultural and structural elements of visual narratives such as comics, manga, manhwa, manhwa, graphic novels, and webtoons. This chapter presented the background, statement of the problem, objectives, and research questions of the study, the significance of the study, and established the theoretical foundation for analysing cross-linguistic variations. These sections explored how the selected visual narratives—*Hada Bhoda*, *Doraemon*, *Patua*, and *Emaki* scrolls could be studied to understand temporal and cultural variations in visual communication. It also discussed how cultural contexts influence visual meaning-making processes.

Chapter 2 provided a comprehensive literature review that contextualised the study within existing theoretical and methodological frameworks. It began with a methodological review highlighting content analysis as the primary approach. The thematic review examined key concepts such as cross-linguistic analysis, McCloud's and Cohn's perspectives on comics and visual narratives, and their emerging views in cognitive science. Theoretical frameworks included Chomsky's theory of innateness, Halliday's systemic functional linguistics, Wittgenstein, Pinker, and the role of Saussure's diachronic, synchronic approaches, and Austin's speech act theory in understanding visual narratives. Finally, the chapter identified research gaps to justify the present study.

Chapter 3 outlined the research methodology, guided by a pragmatic philosophy and a qualitative approach. It detailed the theoretical framework, conceptual design, sampling strategy, and data collection methods used to explore cross-linguistic and cross-cultural differences in Bengali and Japanese visual narratives. The chapter also addressed the analysis techniques applied to explore structural, temporal, and multimodal elements, including speech acts, with the help of the theoretical framework. Finally, it discussed the limitations

and delimitations, framing the scope and boundaries of the study's inquiry into culturally embedded visual storytelling.

Chapter 4 presented the findings related to the first research objective, focusing on a cross-linguistic analysis of visual narratives through their grammar, structure, and social contexts. Using Chomsky's transformational generative linguistics (TGL), the structural and grammatical elements of *Hada Bhoda* and *Doraemon* were analysed. Halliday's systemic functional linguistics (SFL) was employed to analyse functional aspects, revealing how themes like family dynamics, friendships, schooling, sportsmanship, and culinary culture reflect each society's values within their visual narratives.

Chapter 5 addressed the second research objective by analysing the temporal aspects of visual narratives in Bengali *Patua* scrolls and Japanese *Emaki* scroll paintings. The study in this chapter applied Saussure's diachronic and synchronic linguistic theories to explore how time, sequence, and narrative flow are structured within these traditional art forms. By examining panel order and social contexts, the chapter highlighted cultural interpretations of temporality and how each tradition embeds historical and social narratives into visual storytelling across time.

Chapter 6 presented the findings related to the third research objective, focusing on the application of speech act theory to cross-linguistic visual narratives. Using Austin's foundational work and Searle's refined model, the analysis explored how communicative acts, facial expressions, ideophones, and drawing styles function as performative tools in visual storytelling. This chapter highlighted the role of multimodality in conveying meaning and examined how cultural contexts shape the ways in which speech acts are visually represented across Bengali and Japanese narrative traditions with the help of *Hada Bhoda* and *Doraemon*, respectively.

Chapter 7 offered final insights by summarising the key findings and reflecting on the cross-cultural and cross-linguistic variations in visual narratives. It revisited how culture influences narrative structure, temporality, and communicative acts, aligning the findings with the research objectives. The chapter also re-evaluated the theoretical frameworks used and provided concluding thoughts on the implications of the study. Finally, it highlighted the study's contributions and suggested future directions for research in the evolving field of Visual Linguistics.

Chapter 2

Literature Review: Conceptualising the Study

2.1 Methodological Review

Primarily, a content analysis methodology was used to accomplish the goal of this study. According to Neuendorf (2017), content analysis is defined as the objective and systematic quantitative analysis of the characteristics of a message. Krippendorff (2004) rightly pointed out that ‘all reading of text is qualitative even when certain characteristics are later converted into numbers’ (p. 16). Fraenkel and Wallen (2006) defined content analysis as a method used by researchers to examine human behaviour indirectly by analysing their communications.

Although this study does not follow a full-scale ethnographic approach, a very small aspect of ethnographic fieldwork was incorporated on a limited scale to support the second objective. Whitehead (2005) argues that fieldwork is central to ethnography because it enables the researcher to achieve emic validity, which is to understand the cultural system from the perspective of the people within it. During a brief field visit to *Naya, Pingla*, the engagement with *Patua* artists allowed for a closer view of the performative and narrative practices that accompany visual storytelling traditions. While classical ethnography entails prolonged immersion in a host community, according to Wolcott, even short-term observational engagements, when conducted reflexively, can provide valuable cultural insights. These interactions did not constitute a comprehensive ethnography, but they were crucial in contextualising the content analysis of the *Patua* scrolls and ensuring that the analysis remained sensitive to the artists' lived experiences. Whitehead further reminds us that fieldwork in ethnography is not merely about data collection but about building relationships and recognising the inter-subjective nature of knowledge production, which aligns with the qualitative orientation of this research.

Content analysis is usually grounded in empirical data, and it is used to explore topics with predictive or inferential intent. Terms such as ‘sign’, ‘symbol’, ‘significance’ and ‘logic’ have Greek roots. Nevertheless, the ancient Greeks’ interest in language was highly prescriptive and classificatory rather than an empirical one. Aristotelian logic set the benchmark for clear expression, and much of rhetorical theory was directed towards a normative conception of persuasive argumentation (Krippendorff, 2004). This led to restrictions resulting in the artists’ freedom. He further explained that the development of science as an exploratory process, rather than simply declaring conclusions, is a relatively recent accomplishment. George Boole and his peers believed that human cognition followed logical (Boolean) patterns and that human behaviour was entirely rational. Nevertheless, computers built on Boolean logic proved to be disappointing thinking machines. Therefore, through empirical research, modern psychology is replacing Aristotelian frameworks with a ‘psychologic’ approach. In modern times, we no longer measure communication by how well it transmits information, but rather by what happens to the relationships among people as they involve and engage in conversation.

Qualitative content analysis is based on a humanistic tradition that follows an inductive process, unlike quantitative content analysis, which is rooted in a positivist approach. Though not its primary goal, yet, qualitative content analysis may lead to testable hypotheses. It begins with broad, open-ended questions, referred to as ‘foreshadowing questions’, which help in guiding and shaping the research and data collection process, respectively, instead of starting with specific hypotheses. The text takes on a slightly different role in this method; the researcher thoroughly examines the data and, in the process, may discover new patterns and concepts that were not initially anticipated. When this happens, the researcher gets the flexibility to adjust their focus and research questions accordingly, to explore these unexpected, yet significant, patterns (White & Marsh, 2006).

Despite its strengths, qualitative content analysis also has important limitations, especially in the context of cross-cultural visual narratives. While the method allows for deep engagement with the symbolic and structural dimensions of visual texts, it can struggle to fully capture the cultural performativity, oral traditions, and socio-historical contexts that bring these narratives to life. As such, there is a risk of imposing external interpretive frameworks that may not align with how creators or communities themselves make meaning from these texts. Moreover, content analysis tends to prioritise patterns and categories, which may obscure ambiguities, contradictions, or evolving meanings that are central to traditional art forms like the *Patua* or *Emaki* scrolls. These limitations necessitate methodological caution: researchers must remain reflexively aware of the gap between analytical coding and lived experience, particularly when studying culturally embedded art forms that exist in a dynamic relationship with tradition, memory, and identity. This tension informed the methodological decision to integrate limited field-based data into the study.

The primary goal of content analysis is to study social behaviours and experiences without influencing them. This could be accomplished by examining any form of visual or written human communication, such as items appearing in television commercials, novels, newspapers, magazines, television, speeches and many others (Fraenkel & Wallen, 2006). This study predominantly employed qualitative content analysis to examine selected visual narratives—*Hada Bhoda*, *Doraemon*, *Patua*, and *Emaki* scroll paintings. However, recognising the limitations of content analysis in capturing the lived experiences and performative aspects of traditional art forms, the method was complemented by a semi-structured interview and observational fieldwork for a small section of the second objective in this study. This combination enabled a more nuanced interpretation of cultural contexts, especially concerning indigenous narrative practices like the *Patua* tradition. The qualitative factor allowed for the analysis of the readers' voices of the selected visual narratives through the artists' or writers' perceptions; this was analysed using a theoretical framework based on

this study, which is discussed in detail in Chapter 3.

Graphic narratives have an impression of having superheroes as their dominant theme. But the truth is that we are currently experiencing a great revolutionary era in the history of comics and graphic novels. With this, we can bring a whole new dimension to Linguistics studies. It is making the world an interesting place with a lot of imagination and perceptions across cultures. By acknowledging the limitations of relying solely on textual analysis, this study aligns with recent scholarship that emphasises the value of integrating textual methods with ethnographic perspectives (Chute, 2016; Lent, 2001). Comics and visual storytelling forms often blur the boundaries between text, performance, and cultural ritual. Thus, a multi-method approach allowed for a richer understanding of not only what these narratives depict, but also how and why they are created, shared, and interpreted within their respective communities. Such an approach is crucial when dealing with both historically embedded and contemporary visual traditions that transcend purely textual logic. Throughout history and beyond, comic book illustrators or writers have ignited hope in people through the affiliation of the imitation of real-life issues with this humble form of art—if books and films have such great impacts on us and our societies, across cultures, then why not comics?

2.2 Thematic Review

Visual Narratives are a popular medium that is perfectly suited to tell stories. Like an album of pictures that depicts all the moments in a trip, comic strips capture all the essence in a narrative. Speech balloons are dynamically drawn, and the algorithm is fairly simple but good enough (Alves et al., 2008). Comics are generally associated with American superheroes. It first originated in America with the first-ever comic book called ‘The Yellow Kid’, which was first published in 1897 (Diep, 2019). Over time, such visual narratives have been inherited by several cultures. The Bengali comics, *Hada Bhoda* and the Japanese manga, *Doraemon*, along with *Patua* and *Emaki* scroll paintings, were selected as stimuli for a cross-

cultural analysis concerning the study's objectives from linguistic perspectives. This chapter explains several linguistic theories and their applications to visual narratives. The aim is to demonstrate the relationship between linguistic features and visual features through an elaborated theoretical framework, as already mentioned in the previous chapter. This technique will highlight the inherent link between visual narratives and linguistic theories, revealing the underlying nature of Visual Linguistics.

2.2.1 Cross-Linguistic Analysis

Culture and development of human civilisation are deeply interlinked. Culture differs from one society to another, yet there are certain similarities and universalities too; so, cross-cultural research is necessary. The most striking feature of human beings is their diversity. To understand this diversity, we must begin by carefully describing it (P. Spradley, 1980). Most of the diversity in the human species results from the cultures that each human group has created and passed on from one generation to the next. According to P. Spradley, the explanation of cultural differences depends on making cross-cultural comparisons, which in turn depend on adequate ethnographic studies. Much of the comparative work in anthropology has been hampered by shoddy ethnographies, often caused by investigations that impose Western concepts onto non-Western cultures, thereby distorting the results. This view is echoed by Smith (1999), who critiques early ethnographic practices for often projecting Western norms, values, and assumptions onto indigenous cultures. Such impositions, according to Smith, have historically distorted the representation of non-Western societies by framing them through a colonial lens. This not only undermines the validity of ethnographic findings but also perpetuates epistemic injustice, where indigenous knowledge systems are devalued or misrepresented. Smith advocates for a decolonising approach to research that centres indigenous perspectives and methodologies, challenging the dominant Eurocentric frameworks that have traditionally shaped anthropological inquiry. This need to prioritise culturally grounded frameworks in research extends to linguistic analysis as well,

where universal patterns must be interpreted through culturally specific lenses.

According to Wierzbicka (2006), studies using the natural semantic metalanguage (NSM) framework have shown that, universally, every language has a word for 'SAY' that is used in a common structure where one person communicates something to another. This emphasises the fact that it is a shared human behaviour across cultures. However, beyond the basic universality, every language organises its vocabulary for speech in unique ways, giving birth to specific terms for speech practices that are important within that culture. Some speech categories, like 'questions' and 'greetings', appear in many languages, but their meanings often vary. For example, English distinguishes between 'answers' and 'replies', which is not commonly found in other European languages. On the other hand, terms like *demande* in French, *Bitte* in German, and *pros'ba* in Russian maintain a shared European concept, which does not have an exact equivalent in English; the closest term is *request* in this case, which often comes with a formal tone.

Similarly, in visual narratives, we can notice a likely phenomenon to the linguistic differences: certain universal storytelling elements exist, but each culture has its essence and expresses these visual features uniquely. Just as all languages use the word 'SAY', visual narratives too, universally, depict functional concepts through social aspects, which are discussed in detail through the findings of this study. Each culture has its unique visual language that dictates how these functional themes are portrayed and shaped by societal values and art forms.

For example, while Bengali comics and Japanese manga both include dialogue scenes, they differ in the usage of visual elements such as panel structures, expressions, and ideophones to convey emotions. Japanese manga often uses exaggerated facial expressions and visual symbols such as big droplets of sweat or vein-popping marks to signify emotions according to different contexts. Whereas, Bengali comics emphasises more on realism, focusing on

speech bubbles and dialogue tags. These culture-specific conventions make visual narratives complex to translate directly across cultures without losing the true essence. This is the reason why cross-cultural studies in visual narratives can offer valuable insights into the cultural lenses of the picture stories. As a result, these cultural and linguistic differences in categorising speech make cross-linguistic comparison complex and often challenge speakers to understand and explain each culture's unique approach to communication (Wierzbicka, 2006).

According to Cohn (2012), future research on visual language grammar needs to define the principles that guide how graphic sequences unfold and consider the role of semantics and morphology in shaping these principles. Just like in verbal language, where nouns and adjectives have typical semantic associations, visual narratives, too, may have meaning-based constraints on how panels operate within a sequence. For instance, an action symbol might represent an entire event, yet might feel out of place in the first panel, implying that certain semantic features determine its ideal position. Furthermore, cross-cultural studies, like McCloud's (1993) work that compare panel transitions in American, European and Japanese visual narratives, could shed light on how similar aspects of Linguistics can be applied in visual narratives differently across cultures, revealing both universal patterns and culturally specific uses.

2.2.2 McCloud on Comics

In Visual Linguistics, visual narratives or comics are well-known and popular art forms. But unfortunately, people are introduced to this art form in the cheap magazines of childhood or in the funny pages of a newspaper. But nowadays, comics play a major part in documenting the rise of the world's civilisation (McCloud, 1993). Comics is a sequential art form that was coined by Will Eisner (1940). Modern research shows that ancient Egyptians documented in sequential wall paintings for the collection and taxation of crops. 'Beaux Tapestry' created in

France in the 11th century CE, is a huge embroidered wall hanging depicting the history of the Norman conquest of England.

McCloud (1993) focuses on defining comics as a sequential art form categorising single-panel comics as cartoons. The invention of the Printing Press (1439 CE) resulted in the spread of the written word to the masses. With this spread, the humorous and moralistic illustrated stories appeared. In England, Hogarth, one of the famous painters, painted a series of satirical canvases detailing the fall of a woman of easy virtue or the failure of a family's fortune. He made engravings of them to sell to a larger audience. The Printing Press made a vast new reading public. Many artists and writers appeared to exploit the situation by publishing their artworks to the masses.

McCloud (1993) in his 'Understanding Comics' is responsible for introducing the term 'icon' to describe an image representing a real person, place, or thing or idea. Symbols are a sort of icons, as are words and pictures. For instance, one can take a real picture of a face: an oval with two dots for the eyes and a line for the mouth. He calls this icon a cartoon. By doing so, comic artists amplify the meaning in a way that realistic art cannot. The reader can identify comic icons when an artist depicts the icon successfully; it attracts global attention, as we find Hergé's 'Tintin' allows readers to relate to the main character. He further discusses that in comics, the words or script are bolder and more direct. The text must match the iconic pictures; the language is then communicable cross-culturally. It, in fact, seems like something we are born with, and this is what Chomsky's (1975) 'innate ability' hypothesis explains. Referencing Freud, Bender (1944) writes that comics use a kind of language that resembles a child's early attempts at mastering language. This is the very stage of children when they start playing with sounds and words, and find connections between their thoughts. Comics are highly flexible in format as well as in content, with a reflection on personal, social and cultural issues. Comics allow writers and artists to experiment with visuals and language, respectively, in ways that defy traditional structures and, as a result, challenge readers to

engage actively with both images and texts, conveying meaning. Therefore, comics or other visual narratives serve as a bridge between inborn language capacities, as per Chomsky's theory of language acquisition, and creative expression, allowing readers to process complex meanings through the blend of texts and visuals.

We can convey emotions through language; comic artists make emotions visible. They include motion lines, stars to show pain, and beads of sweat. It creates the language of comics. McCloud (1993) stresses that in comics, words and images work together. He says that there is no formula for a comic artist to follow. It is more alchemy than science. Comics is a work of art, as other artworks are. But it is fluid with ever-changing activity. 'Maus', a comic book by Spiegelman, won the Pulitzer Prize's special award in Letters in 1992. The story is of the Holocaust, where the characters are animals. Jews are mice, whereas Germans are cats. It was the first graphic novel to be honoured.

2.2.3 Cohn's Views on Visual Narratives

Visual narrative is an emerging field in cognitive science as observed by Cohn and Magliano (2019). These visionary narratives are formed among our oldest records of human intelligence, appearing on cave paintings, wall carvings, and ancient pottery; they pervade across cultures from instruction manuals to comics. Yet, despite this fundamental place in human expressions, the study of visual narratives has only recently gained traction in cognitive science. It has already been seen that visual narratives are sequences of images with meaningful intent, illustrating typically a continuous event sequence to tell a story. Drawn visual narratives appear in comics, picture stories and storyboards. Cohn and Magliano question why visual narratives are not studied with the type of seriousness afforded by other types of human expressions and communications. We see several potential factors involved in visual narratives:

First, sequential images are frequently used in tasks and stimuli to investigate other aspects of cognition. This means the researchers are certainly aware of visual narratives as potentially effective communicative tools. But they just have not consolidated around studying them.

Second, the use of visual narratives as a stimulus presumes that their understanding is fairly transparent.

Third, though attitudes have been changing in recent years, visual narratives have historically been afforded low esteem in culture, in the form of comics or illustrated picture books, compared to language or film.

A growing number of studies have begun to examine the actual properties of the sequential unique understanding (Cohn, 2013). This work has pointed to complex structures that afford questions about visual narrative processing as a unique focus of scientific inquiry connected to fundamental issues in the broader study of cognition (Cohn & Magliano, 2019).

The ability to draw is a basic human ability, and sequential images are prevalent across human history and culture. The fundamental aspect of human expression yields several important questions, such as how such individual and sequential images are structured and how people process and comprehend them (Cohn, 2018). A language constitutes a set of patterns in a person's mind/brain. There is no universal language, but rather diverse languages manifest across the world. The basic idea holds both for spoken and sign languages, but also for visual languages. Not all drawings are visual languages.

A language requires an interaction between primary components: meaning, modality and grammar (Cohn, 2013). Spoken languages express meaning using the modality of phonology (sound), while visual languages use graphic structures (drawn lines). Systematic mappings between a modality and meanings create a stored lexicon. These expressions must be ordered using a grammar. In visual languages, the grammar organises meaningful images with coherent sequences. For our purpose, visual languages appear prevalently in what are broadly

labelled as comics around the world. Visual languages arise in several different cultural contexts. It is important to emphasise the separation between visual language and comics (Cohn, 2018).

Different types of comics may be characteristic of different visual languages. Japanese language is highly associated with manga, while Kirby American visual language characterises superhero comics from the United States- i.e., the drawing and storytelling styles (Cohn, 2013). Ultimately, the definition of comics rests with a web of socio-cultural ideas, including their cultural contexts, genres, and possibly the different visual languages they use (Cohn, 2018).

2.3 Theoretical Review

According to Scovel (1998), when we are babies, we cry for food and eventually for other needs; gradually, we start talking, and as a result of it, we start acquiring languages. Scovel also discusses ‘linguistic thinking’, which, according to my knowledge, is basically the visual component of the conceptualisation of language that is untimely manifested in the gestures which are accompanied by speech. So, the question that arises here is—which one is truly our innate ability—language or visuality or both? In fact, it seems, Visual Linguistics is something that looks like an inbuilt system present in our brain. Every child, irrespective of cultural background, scribbles and gradually grows up into imitating the surroundings through art. It depends on how we see our social environment and relate it to the imitations present in different artworks. It totally takes place in cognition, and when it is combined with language, is it there that it becomes a complete form of comprehension?

Language is a complex human behaviour; neuroplasticity requires a young child’s brain to allocate different areas of the cortex for language. Neuroplasticity and vision go hand in hand: the visual system is not fully developed at birth; over the years, visual experiences form the neural architecture of the visual system; it is also believed that between half and two-

thirds of the brain is used for visual processing (Kolb & Gibb, 2011). So, if languages have a universal grammar, then visuality too should have universality; and there should be a universal way of viewing the artworks present in graphic narratives. Visual meaning is situated within the broad field of social semiotics, which draws inspiration from Halliday's systemic functional theory of language. An idea of every text (here, different forms of narrative structures of picture stories) fulfils a threefold purpose: representing the material, communicative interaction with others' assertions, and linking, referring, foregrounding, and backgrounding (Halliday, 1961).

As the Chomskian theory believes in the innate ability theory of language, it is to be put into consideration that visuality too is about the depths of the mind, about revealing the so-called 'false reality'. Visuality links the mind with the external world; therefore, we can comprehend dreams even when dreams do not contain any language; it is to be noted that this visuality is not barred to what we see with our eyes open (with the help of our retina); blind people can also dream which gives us an even stronger opinion that Visual Linguistics could be an innate ability of human beings. However, the visual components of people who are blind since birth cannot be formed from visual memories or the associated circuitries (Baird, 2020). Therefore, it appears that Visual Linguistics, too, has a universal grammar embedded in each human being that helps us to comprehend graphic stories, cross-culturally. An argument supports the claim that language is innate because children do not always utter words and sentences that they have heard other people utter. If language were not about its innateness, then it should be expected that our children will comprehend only those particular sentences and words that they have heard others utter (Cummins & Davidson, 2007). This is again true for Visual Linguistics, as it comes with several forms of imagination and creativity in children. The Chomskian theory explains that all children, regardless of their cultural backgrounds, share similar internal challenges, which differentiate slightly the grammar that they end up constructing (Chomsky, 1977, p. 98). It should be the same with children from

different cultural backgrounds when they perceive the narrative structures through their linearity, movements, and production of shapes, colours, and vectors in picture story books. Lenneberg (1967) states that language materialises before it is even required. For example, among children, the emergence of language takes place when they are between 12 and 24 months of age. At this point, the child relies entirely on its parents for survival. On the other hand, visuality is something that plays a role right after birth, even without relying on parents. The enigma here is—visuality plays a very primary role in ability, which adds up to the whole supposition of the theory of innateness.

2.3.1 The Chomskian Hypothesis: The Innate Module of Language

There is an idea that the human brain is organised into different modules and each with different functions: this is exactly what the Chomskian theory of innateness explains. The very influential work of Chomsky (1957) was ‘Syntactic Structures’, which talks about Zellig Harris’s model of transformational generative grammar. It contains the ever-so-famous and loved by all linguists, the sentence, ‘Colorless green ideas sleep furiously.’ Chomsky used this sentence as a grammatically correct one with no proper meaning to it. Therefore, Chomsky argued that irrespective of semantic values (the study of meaning), there is a syntactic value (the study of sentence structures); this syntactic structure is embedded in our mind, mental processing and cognition. With this notion, it is theorised by Chomsky (1957) that human beings acquire language rather than learning it because of the very presence of this syntactic structure in the human brain.

This conception of innateness is extensively used in various debates in the biological and cognitive sciences. He argues that the speed and reliability with which children acquire the ability to understand and speak a human language with the level of syntax, the impoverished nature of the linguistic stimulus that children receive, and the small impact that the variation in the stimuli seems to have variations in language acquisition, indicate that children do not

acquire their knowledge of language through learning; so, it must be innate (Mameli & Bateson, 2011). Nevertheless, no one has yet been able to provide an account of genetic encoding that can be used to explain what innateness actually stands for.

According to Skinner (1957), children learn language when their verbal systems are brought under the control of environmental conditions as a result of training by their caregivers. The children are approved and punished accordingly for their various linguistic productions, and as a result of it, their dispositions to verbal behaviour gradually converge on those of the wider language community. Therefore, one understands the utterance, 'Shut the door,' to the extent that one responds to it.

It has generally been observed that children do not need any kind of formal teaching or training in learning to speak. Although Skinner (1957) thought otherwise. It is the natural outcome of a child to speak by listening to their parents and others. Chomsky (1957) believed that children do not need any kind of formal teaching to learn to speak. The human brain is structured in such a way that they are born with a natural capacity to learn language. The brain contains certain systems for recognising patterns of sounds. He further explains with his hypothesis that all human beings are born with an acute knowledge of language and language structure; this language is used in the course of language acquisition. He believed that there is a critical time period for language learning. Nevertheless, every human being is born with innate principles of language. Children usually learn language spontaneously and speak creatively. To support this hypothesis, he has looked into the facts of the complexity of human language systems, the universality of language acquisition, the facility that children demonstrate in acquiring these systems, and the comparative performance of adults in attempting the same task. The critical period of this hypothesis states that the first few years of life are the crucial time for someone to acquire a first language if presented with adequate stimuli. If the language input does not occur until after this time, then the individual will

never achieve a full command of language, especially grammatical systems. Therefore, Chomsky (1957) effectively demolishes this theory of Skinner (1957) of both language mastery and language learning. He argued that mastery of a language is not merely a matter of having anyone's verbal behaviours controlled by several elements of the environment, inclusive of others' utterances. He further argues that the use of language is stimulus-independent: any word can be spoken verbally in response to any environmental stimulus, depending on one's state of mind. It is also historically unbound: our utterances are not determined by our history of reinforcement, as it is clear that we can utter things that we have not been trained to say.

So, it appears that a child's brain is very active in mapping. Different languages have different structures. If a dog barks in a Bengali family, it will still bark in a French family. But if a Bengali child is adopted by a French family right at birth, then this child will not speak the SOV structure of the Bengali language; rather, it will acquire the SVO structure of the French language. Does it work the same way with reading different structures of visual narratives? Unlike spoken or written languages with distinct structures, the layout of a visual narrative, which reflects its structure, does not develop in the same manner as language. Humans appear to have a natural ability to produce mental films, which allows us to understand our dream sequences and form stories without relying on language. This begs the question of how, despite their various structures, visual narratives share parallels with language qualities in establishing a sequential knowledge of the storyline across cultures.

Chomsky (1957) stated that grammar must be enriched with a second type of rule, known as 'transformations.' The early vision of syntactic theory has evolved ever since he made it. What has not been changed since then is the rules governing the syntactic structure of sentences and phrases, stated in terms of syntactic categories that are highly abstracted from the properties of utterances that are accessible to experience. This works similarly in the case

of visual narratives or comics. Since the 1980s, several scholars have relied on the notion that comics are a language. Both Eisner and McCloud (1993) have written about ‘the language of comics’, and this metaphor has given some common ground for discussing art and research (Bramlett, 2012). The idea of visual narratives as language is interesting and appealing because sequencing in language aligns very well with the notion of sequencing in comics. However, from the linguistic point of view, comics are not or cannot be language. Rather, comics or visual narratives rely on visual grammar—a set of conventions for sequencing images and texts, using spatial arrangements, and conveying emotions and actions through art and ideophones. Therefore, it has linguistic properties even though it is not a language on its own. This notion again draws the attention back to the sequential way of understanding language (or understanding the picture sequences) and that it is the innate ability of language, and if things and events possess any intrinsic true essence, which could produce an important space for cross-cultural variations affecting individual cognitions, while considering the theory of innateness.

2.3.2 Halliday’s Systemic Functional Linguistics: Its Relation to Chomskian Theory and Visual Linguistics

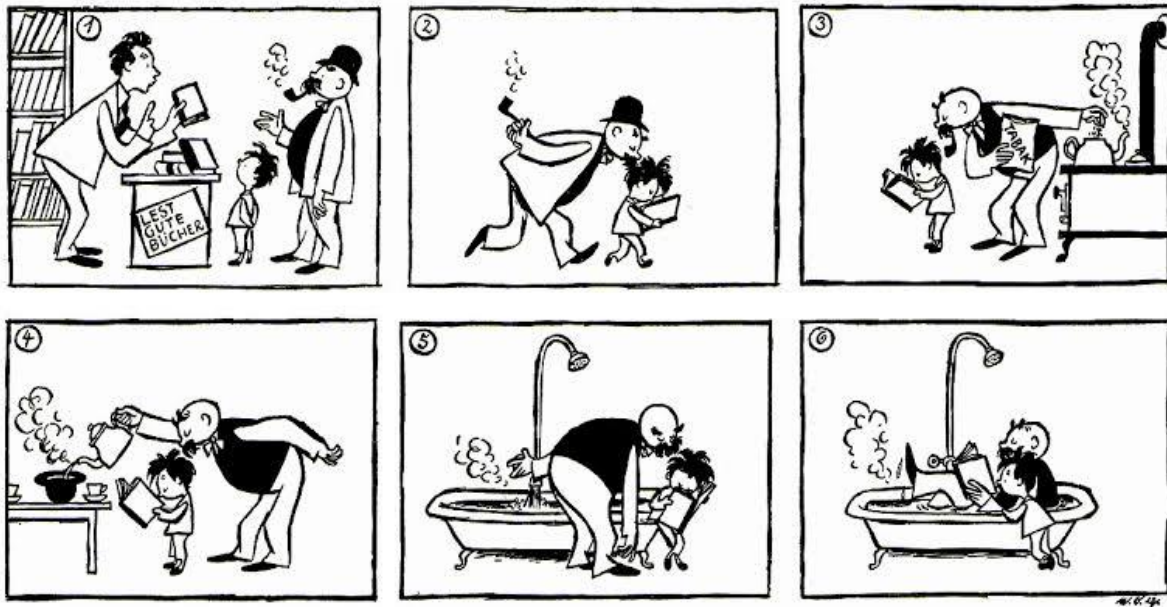
Linguistics has witnessed a considerable number of theories, and each one has its distinct orientations, trends, and subjects of study. Yet, Halliday and Chomsky’s traditions have been initiated by a large number of followers. Chomsky’s (1957) transformational generative linguistics (TGL) and Halliday’s (1961) systemic functional linguistics (SFL) have been deemed as two of the most influential and pivotal traditions in the linguistic academic field. Though both systems are distinct in many respects, making the same impacts, there are a number of differences between these two systems. For example, systemic functional grammar (SFG) studies the language through meaning (i.e., its function). While transformational generative grammar (TGG) is a fully influential and developed version of Linguistics through form. Halliday (1961) theorised that Linguistics should describe actual sentences with many

functions and without a deep structure. In addition, his theory is concerned with the function of the sentences, or in other words, the writer's purpose in writing the sentence (Almurashi, 2016).

On the other hand, Chomsky (1957) maintained that Linguistics should go beyond merely describing syntactic structures, and his purpose is to elucidate why language is structured in the way that it is. To summarise, Chomsky (1957) characterised form independently of function and meaning, while Halliday (1961) had good reasons to believe that functional meaning can help shape form (Bavali & Sadighi, 2008).

In the case of visual narratives, this study believes in both structural and functional aspects. Without structure, we would not be able to comprehend the panels present in the visual narratives, be it comics, graphic novels, manhwa, manhua, manga, or recent times' webtoons: these diverse forms of visual narratives present in various cultures combine art and text generally in sequential manners to convey stories, ideas and emotions. The comprehension of sequential images draws upon a narrative structure that is organised into constituents, analogous to the grammatical structure of language (Cohn, 2014). If the panels are not in order or in a sequential manner, the function would get lost:

Figure 2. Picture Story: The Captivating Book



Source: Plauen (2003)

Here, the structure of comics (establisher, initial, peak and release) as established by Cohn (2012) conveys the function and gives the readers the whole story:

- 1) The comic strip starts with an **establisher**, where the father is buying a book together with his son; ➡
- 2) In its **initial** stage, the father runs behind the son with his pipe, reading the book with him; ➡
- 3) In the third panel, which is the **prolonged initial** stage, the father keeps on reading with his son while making tea; ➡
- 4) It reaches its **peak** when the father starts pouring the tea unmindfully into his hat instead of the cup while he still reads the book, thus invoking our laughter; ➡
- 5) The **peak prolongs** as the father enters the washroom, dropping the 'hat of tea' in the bathtub while he is still reading the book along with his son; ➡

- 6) The father finally, himself, gets into the bathtub instead of his son and takes the book in his hand while his son accompanies him in reading, considering a **release** to the tension and thereby coming to a conclusion.

Therefore, Chomsky's TGL (1957) and Halliday's SFL (1961) are both important in visual narratives for constructing a story based on structure. One of the most distinctive human behaviours is the ability to create mental stories. This raises the question of how our brain's natural ability enables us to comprehend visual narratives sequentially. Moreover, if events and things carry intrinsic meanings, this would imply significant cross-cultural variations that impact individual cognition, aligning with the theory of innateness. This concept leads to the perplexing observation that artists from the same cultural background can create vastly different works, yet people across different cultures and times can perceive and interpret these artworks in similar ways. In this study's analysis of visual narratives, Chomsky's TGL (1957) and Halliday's SFL (1961) are intertwined, with structure and function complementing each other to facilitate comprehension of these picture stories.

This concept leads to the perplexing finding that artists from the same cultural background might create completely diverse works, while others from different cultures and times receive and interpret these artworks in comparable ways. In my analysis of visual narratives, Halliday's SFL (1961) and Chomsky's TGL (1957) are intertwined, with structure and function complementing each other to facilitate understanding of picture stories.

It has already been seen that Halliday insisted that the central concern of Linguistics should be the study of language through meaning, which is different from the dominant approach of Bloomfield (1933). He rejected the possibility that Linguistics analyses meaning. He was fully convinced of the need for linguists to study oral language instead of studying written documents (Almurashi, 2016). According to him, the documents do not fully represent a spoken language due to the fact that the language undergoes changes over time, and what

something means today might have meant something different in the past.

When it comes to visual narratives, what Bloomfield (1933) believes does not seem to be true, and at the same time, it seems to be quite true. A picture is visual as it is not oral, and the narrative is written. Of course, in ancient traditions, stories were often orally sung by performers along with certain displays of scroll paintings (Ahire, 2020). But in picture stories nowadays, it is documented with pictures and written texts.

However, in SFL, a text is analysed in four ways: context, semantics, lexico-grammar and phonology. Halliday models the concept of a situation where the aspects of the context relate intimately to the language used to create text, in terms of three important strands (Matthiessen & Halliday, 1997). They are as follows:

- a. Field:** It gives us an indication of the topic or what is being talked about.
- b. Tenor:** It gives us an indication of who is/are involved in the communication and the relationship between them.
- c. Mode:** It gives us an indication of what part the language is playing in the interaction and what form it takes (written or spoken).

Thus, SFL is characterised as an applicable linguistic theory, which means it is designed to have the potential to be applied to solve the problems arising in communities around the world (Matthiessen & Halliday, 1997). Halliday (1961) writes that he aims to make a coherent tradition of language that is applicable in the sense that it can be beneficial to the majority of people engaging with language in the course of their work. In visual narratives or picture stories, it seems that it is quite applicable: language plays an important role side by side with pictures, moving sequentially through the panels. SFL is concerned with attempting to comprehend the manner by which language is utilised for various reasons and in various scenarios, thereby serving as a communicative motivation in the learning of language

(Almurashi, 2016). Thereby, it can be said that a child learns to communicate with sequential pictures through writing or vice versa. To sum up, SFL, developed by Halliday considered a chief force in World Linguistics. It is also considered an influential tradition in linguistic study as it reflects the trend in linguistic development.

2.3.3 Wittgenstein and Pinker in Visual Linguistics

Wittgenstein (1953), on the other hand, emphasises the individual use of words in a language or the names of the objects in the form of sentences which are combinations of such words or names in his 'Philosophical Investigations'. In his 'picture of language', it is found that the roots lie with the idea that 'Every word has a meaning. This meaning is correlated with the word. It is the object for which the word stands.' It means there is a relationship between words and objects through language. We may say that in the case of words and objects, the relation is of meaning—the meaning of being an object for which the word stands. The meaning seems to depend on the meanings of the words making up the sentence (Kumar, 2019). Wittgenstein gives an example where a man asks for five red apples in a shop. The shop owner draws a drawer marked 'apple'. Then he opens a coloured chart to match the colour red; after that, he counts for each apple from one to five. Thus, it is the use of the words in a sentence that makes the meaning. In his 'Philosophical Investigations', he aims to describe how words get their meanings. In picture stories, the readers thus use their cognitions to relate the interlinked sequential relation of the words and their relations with the pictures.

In his later days, Wittgenstein seemed to reject his earlier notions of the uniformity of language, which would restrict the word to a rigid and demarcated use—a use which would suit all cases. He came to think then that language is flexible, subtle and multiform. He makes use of words like 'bricks' and 'slabs' which refer to objects, and *prima facie* it looks like logical atomism with the concept that the elements of language reflect data in reality (Ara,

2006). When a builder utters the word 'slab', his assistant brings him the object that is needed. So, it needs training to understand that, on hearing a certain task, just as a child learns the use of words on the basis of demonstrative teaching. A kind of rapport is to be established between the builder and the assistant. Likewise, in visual narratives, there has to be an agreement in one's reactions, which makes the sequential way of understanding comics comprehensible. Wittgenstein (1953) wants us to see that natural language use is pivotal to understanding the nature of the mind. In 'Philosophical Investigations', he is primarily concerned with human behaviour. For him, the language game is a communal process: It is language in action, language as behaviour, a form of life (Byrne, 2012).

Wittgenstein (1953) now sees language as a tool rather than his previous description in 'The Tractatus' (1921), where he described language as a picture. He thus negates his earlier observation that language is a picture. He wants us to think of words as tools and sentences as instruments (Byrne, 2012). So, in a picture story, a reader is supposed to view the words and sentences as tools and instruments to appreciate the given picture; when there is no text in picture stories, the reader would only be a viewer to appreciate the pictures and create a story sequentially in their mind.

Wittgenstein (1953) focuses on the flexibility and context-dependent nature of language that aligns closely with visual language, where meaning is derived from images as well as texts through the interplay within a culture-specific context. Just as he argues that the meaning of a word arises from its usage in particular settings, the interpretation of visual elements is similarly influenced by the cultural conventions and shared experiences of the readers. In visual narratives, the combination of images and texts creates a complex language system where each visual and textual feature contributes to the overall meaning. This reflects the Wittgensteinian notion of language as a tool in the context of picture stories, where all the visual elements, along with texts, serve as instruments to guide the readers' cognitions and

comprehension. The readers, therefore, engage actively with all the elements present in visual narratives, drawing upon their cultural backgrounds and cognitive frameworks to construct meaning. Moreover, the idea of the ‘language game’, where understanding is rooted in communal practices, can be extended to visual narratives across cultures. Different cultures write their visual narratives through different visual grammar, symbols, and narrative structures that will be discussed thoroughly in the findings of this dissertation. Furthermore, Wittgenstein’s (1953) insights provide a framework for understanding how visual narratives communicate meaning. By recognising the unique relationship between texts, images and cultural contexts, the complexity of visual storytelling and how it reflects the multifaceted nature of human experience can be understood.

On the other hand, Pinker (1994), in his first popular science book, ‘The Language Instinct’, made the case for an innate language capacity: language is not learnt in its entirety, much of it (particularly grammar) is already present in a determined structure of our mind/brain. Much like Chomsky (1957), he establishes the innateness of language learning. In his ‘How the Mind Works’ (1997), he divided the mental abilities into independent modules. For example, the process of dealing with language is not related to the process in other domains, like numeric or spatial cognition. These modules are computational adaptations shaped by evolution. We all share them as it is built into the brain.

So, Pinker is an evolutionary psychologist and linguist who believes that much of what defines us is built into the structure of the mind or the brain (Meteyard, 2009). Pinker scientifically presents his side of rationalism and modern Darwinian psychology. To him, language is a digital medium that uses features of meaning (such as agency, causation or spatial relation). These features show how language curves the world in a way that reflects human experience.

In this way, a verb can take in different constructions expressing similar events in different ways:

a. I **broke** his heart.

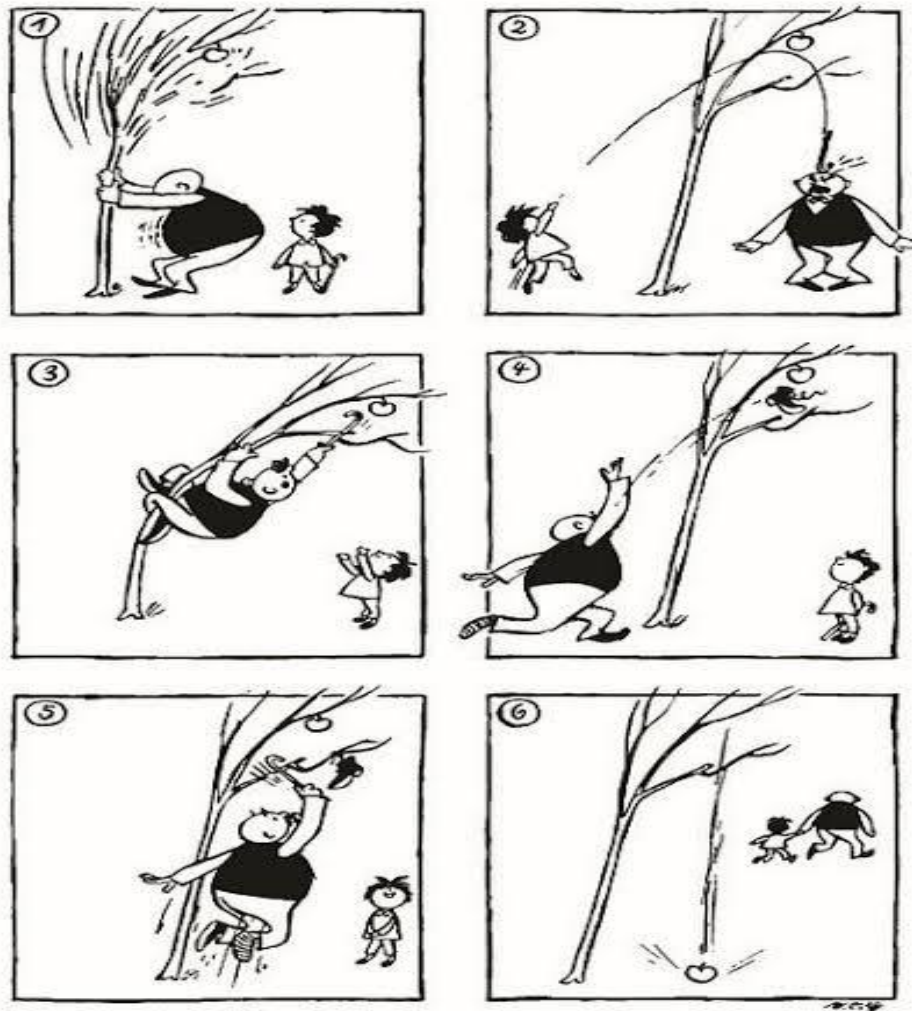
b. The heart **broke**.

Here, 'break>broke' is the action which can be caused by someone, or it can simply happen.

In the first sentence, there is an agent who is causing the heart to break; whereas, in the second sentence, the heart is simply broken.

Similarly, this could be explained with the help of the following visual narrative:

Figure 3. Picture Story: The Last Apple



Source: Plauen (2017)

Here, there are two agents in the picture story trying to make the apple fall, but it did not: being exhausted, the father and the son left. Just after they left, the apple fell without an agent.

Here, the verb 'fall' is in our mind, and readers across the globe can comprehend the action. Pinker (1994) made the case for an innate language capacity. The verb 'fall' is already present in the determined structure of our mind/brain. So, the reader could spontaneously

communicate through the visual narrative. Here lies my quest about the important space for cross-cultural variations that could hold any intrinsic true essence, considering the theory of innateness through the sequential manner of understanding visual narratives.

2.3.4 Diachronic and Synchronic Linguistics in Visual Narratives

Strings of pictorial representations—also explored within Visual Linguistics—possess a grammar of their own, much like verbal language, as previously noted. Meanings are conveyed through sequences of images. Quite interestingly, experiments have not yet been done in online comprehension of static visual narratives like comics; contrarily, with films, it is found that the viewers can consciously identify the semantic shifts between individual film shots (Cohn, 2014). In a situation like this, what can be argued for is the following speculation: the semantic shifts, which are often noticed in the case of processing the images ordered in sequences, could be established or identified by the interpreters (=viewers/readers) because of some underlying principles of visual cognition interacting with the universally shared linguistic substratum. This concept aligns with the principles of diachronic linguistics, where shifts in meaning or interpretation are examined over time, similar to how semantic changes in visual sequences can evolve as they interact with underlying cognitive and linguistic structures across different periods of interpretation.

Diachronic linguistics is the study of a language through various historical periods. One of the two primary temporal dimensions of language study described by Saussure (1916) in his ‘Generative Linguistics’ is diachronic linguistics (Nordquist, 2019). On the other hand, synchronic linguistics is the other; it is the study of a language at a specific period (usually the present). The terms diachrony and synchrony refer to an evolutionary phase of language to a language state, respectively. For example, analysing the word order in a sentence in Old English would be a study of synchronic linguistics. Diachronic research, on the other hand, would look at how word order changed in a sentence from Old English to Middle English to

present English. Altogether, the terms diachrony and synchrony refer to an evolutionary phase of language to a language state, respectively.

In the case of visual narratives, this entails examining how language performs within the context of a specific panel, page, or visual. Understanding the present environment is essential for understanding linguistic nuances. If we take scroll paintings as the father of comics or manga, then it could be considered a well-established reference to the evolution of visual narratives. In scroll paintings, we often find that inscriptions precede or follow the image. These inscriptions may contain a poem or a note written by the painter that explains the events that led to the creation of the painting (Florek, 2014). These scroll paintings were found both in West Bengal and Japan. The scrolls were often folded up, and the *Patuas* or *Chitrakars* would travel from village to village, and in front of a suitable gathering, they would unfold the scrolls and narrate the stories painted on them (Gangopadhyay, 2018). It is also believed that the handscroll culture was invented in India around the fourth century B.C., when it was predominantly used for religious texts and spread to China by the first century A.D. The handscroll culture was later introduced to Japan some centuries later, as part of the spread of Buddhism from the mainland in the sixth century, along with many other cultural advances such as the Chinese writing system (Willmann, 2012).

Despite being universal, it is also found in addition to what is already argued above—certain aspects of the visual narratives differ from one culture to another, as is also the case in language. More specifically, probably, one could even argue that the way we visualise is often influenced by cultural practices, the consequence of which can be clearly noticed in the relative richness of the representation. This might be the reason Mirzoeff (1999) argues that an image attracts much more than reality.

Due to the differences and culture-specific components of visual language, we find a wide range of diversity in the world of visual narratives. Following Watson (1913),

representational constituents of comics, as the overt form of visual behaviour susceptible to cross-cultural contingencies, seek in-depth investigation to understand the way underlying principles of human cognition are manifested in real life. People, in general, perceive the world more or less in the same way, yet we find the differences in drawing those perceptions. Again, individuals in one culture draw more or less in the same manner, but the same people draw differently from individuals in other cultures (Cohn, 2012).

One of the major requirements of any kind of modern art is the unity of both its ideological and artistic aspects. One cannot imagine high ideological principles without high artistic merits. However, this unity is not always seen in artistic practices (Propp, 2009). In the domain of visual narratives, this is reflected in the lack of understanding of the specific rules of the narratives and, thus, in the inability to use them.

2.3.5 Speech Act Theory in Visual Narratives

Speech act theory is arguably the most popular among general theories of language usage (Levinson, 1983). Psychologists suggest that the acquisition of the concepts underlying speech acts may be a prerequisite for the acquisition of language in general (Bruner, 1975 & Bates, 1976). Literary critics have turned to speech act theory for illumination of textual subtleties or an understanding of the nature of literary genres (Ohmann, 1971 & Levin, 1976). In this theory, anthropologists found some account of the nature of magical spells and rituals in general (Tambiah, 1968), whereas philosophers have seen potential application to the status of ethical statements (Searle, 1969). According to Levinson, linguists, on the other hand, see the notions of speech act theory as variously applicable to the problems in syntax, semantics, second language learning and elsewhere. In linguistic pragmatics, speech acts remain along with presuppositions and implicature, in particular, one of the central phenomena that any general pragmatic theory should account for.

There is a vast literature on the subject, and in this study, all the relevant works cannot be reviewed. However, what is attempted here is a brief sketch of the philosophical origin of the speech act theory and a few related to visual narratives. Speech act theory is a subfield of pragmatics that studies how words are used, not only for presenting information but also for carrying out actions. J. L. Austin, an Oxford philosopher, first proposed this speech act theory in 1959, and J. R. Searle, an American philosopher, expanded on it in 1969. In the early 1930s, the new theory of logical positivism rose to prominence (Levinson, 1983). According to logical positivism, a sentence can be verified (tested) for truth or falsity; truth and falsity have long been key topics of study. Wittgenstein (1953) vigorously attacked in his 'Philosophical Investigations' with the famous slogan 'meaning in use' and emphasised that utterances are only explicable about the activities or language games in which they participate. During this particular period, Austin (1959) proposed his theory of speech acts. Wittgenstein's emphasis on language usage is strikingly similar to Austin's assertion that 'the total speech act in the total speech situation is the only actual phenomenon which, in the last resort, we are engaged in elucidating' (1962). Austin expanded on Wittgenstein's discoveries by emphasising the performative aspect of language. He stated that many utterances serve a performative role, meaning that they not only describe or report on something but also perform an action or achieve a goal. For example:

- a. I sentence you to ten years of hard labour.
- b. I give my word.

These sentences are not just statements of intention but also commitments to perform (Levinson, 1983).

Searle (1969) built on Austin's work and developed a systematic theory of speech acts. He argued that utterances consist of three dimensions of meaning: the propositional content, the illocutionary force (speaker's intended meaning or function), and the perlocutionary effect

(actual effect on the listener). He also introduced the concept of speech act types such as requests, commands, promises and apologies, which are universal across languages and cultures. In general, the philosophical background of speech act theory could be seen as a shift away from traditional views of language as a static system of meanings towards a more dynamic and contextual understanding of language (Levinson, 1983).

People are familiar with speech acts when studying pragmatics. Human beings often express their ideas through utterances. There are various types of utterances, including directive utterances (Anwar, 2012). Directive utterances are speech acts used by speakers to persuade one or two others to do something. There are several types of directive utterances, including commanding, ordering, requesting, proposing, and banning. In general, directive utterances can be found in everyday speech or in literary works such as novels, short stories, films, and comic books.

Speech acts in comic books or other visual narratives include various linguistic and visual features that are used to convey information, express emotions, and propel the story along. Comics contain both textual and visual narration, which enables multiple types of narration to occur inside a single comic panel (Tang, 2016). Textual and visual narration kinds collaborate to tell a story in comics. An outward narrative point of view in images is often combined with an inside story. This gives us a clear, invisible camera view of a scenario. Tang argues that a visual external view of a character's actions can be paired with a textual internal insight into the character's mind to provide both external and internal narrative perspectives in tandem.

Though McCloud (1993) claims that comics are composed of 'images in deliberate sequence', his ensuing explanation and examples reveal that he sees narrative as the primary sequential organising element (Pratt, 2009). Characters cannot see speech balloons, but they can hear the words inside, and each character is likely aware of the contents of their

own thought balloons. Pratt claims that when there are sound effects, characters can hear them, although the sounds heard within the diegesis may not be identical to the sounds represented in words.

Speech acts in visual narratives can be conveyed not only verbally and contextually but also culturally, along with visual cues. Characters use facial expressions, gestures, and body language to indicate their intentions, feelings, and attitudes, essentially conducting speech acts with the bimodal usage of texts and images. The arrangements of visual elements within the narrative context affect the comprehension of speaking acts, as viewers derive meaning from image sequences, scene framings and cultural contexts.

2.4 Review of Previous Research

No human cognition or perception is the same. People also differ in how they process information about things or other people. Need for cognition refers to the habit of thinking carefully and fully about our experiences including the social environment or situations that we encounter (Cacioppo & Petty, 1982); it seems that they are all unique exactly like the patterns on the bodies of zebras; cross-cultural perceptions play a major role along with how individualistic ideas could affect it from culture to culture or within one culture.

According to Kramnick's (2018) argument, 'the eighteenth-century theory of ideas and the twentieth-century theory of concepts both understand cognition to work on symbols and symbols to stand in for, or be about, external objects or events.' As a result, cognition depends on signs and symbols that are present in the lexicons of our brain, and with the help of these symbols, we comprehend the world around us. When it comes to certain visual narratives, we comprehend in the same way and try to fit them according to our environments or understandings. This, sometimes, can give rise to major cross-cultural differences or not at all because of the theory of innateness. He further explains that poetry cannot be explained by botany or physics, because it is not physical. Reading is not similar to seeing. Reading has

something to do with the consciousness of each human being that is somehow not the same in each individual, exactly the way we do not have the same imaginations and behaviours. When it comes to visual narratives, it should definitely work the same way because different people might comprehend the narratives of the picture stories differently, and when there seems to be a cultural stigma within the stories narrated, comprehension may have broader differences.

Thiher (1984) states that there is always a connection among the things present in the world, and they do not exist apart from the language that states them. The number is indefinite when it comes to the possible linguistic constructions that might supply these connections. Words also have meanings independent of how they are drawn. When we combine words and images in the bimodal form, the meaning of the set of words relates to the image they are part of. Sometimes the words are exact copies of what is drawn beside them, and sometimes, they appear vague or contrast with each other. Sarcasm, metaphors, ironies, paradoxes, and onomatopoeias are parts of literature. They can also be seen when it comes to visual narratives. As a result, cross-cultural differences appear in perceiving picture stories when the narratives are read or sometimes viewed.

Even though drawing comics or other forms of visual narratives seems more or less similar all over the world, they manifest differently in Western and Eastern cultures. Some visual narratives are so deeply rooted in specific cultures that readers or viewers from other cultural backgrounds may struggle to comprehend the context of the story. This difficulty can often come from political or religious themes, but it is most commonly caused by the idioms and ideophones used, which can lead to major confusion in cross-cultural comprehension. As a result of it, perceptions differ. According to Gigerenzer and Murray, as cited in Danziger (1988), 'There is a final parallel between visual perception and scientific practice concerning the certainty with which both present their products. We experience what we see as certain:

We see that the moon is larger on the horizon than up in the sky; we see the Swiss mountains arising immediately behind Lake Constance when the view is very clear, but under normal viewing conditions, we see them far away.’

When we look at a comic book, we are not seeing the world as it is, nor is it shown in a plain and simple manner. Instead, we are watching how an artist interprets or transforms reality. This includes components that are frequently exaggerated, altered, or wholly created to portray specific ideas, feelings or messages (Wolk, 2007, p.20). This unique combination of visuals and narratives in comics allows for creative expression that goes beyond the duplication of reality, providing readers and viewers with a fresh and imaginative perspective. Picture stories are narratives that come in visual forms, and when we read them, we tend to believe that they are real on some level. Wolk further states that comics are always what they appear to be, and we do not need metaphors to understand the simplistic form of art with some obvious texts. But it is the human brain that makes it unique in its comprehension, respectively, and therefore perceptions take birth. No two people would experience the world in the same way, and similarly, no two cartoonists would draw in the same way, and the way they draw would be the closest to the readers’ perceptions that they experience through their individual eyes.

According to McCloud (1993), readers of comics experience various characters, ideas, events, and emotions of the stories directly when they read them. As a result of it, comics acts as an intermediary between the storyteller and the audience. But this also does not mean that everyone experiences the sequential stories the same way, or that they exactly understand the storyteller’s point of view. Perceptions are unique in every human being. Artists of comics have come up with new techniques and forms of drawing comics since time immemorial. They all have their own ways of telling their stories, and on the other hand, readers have their own ways of perceiving them.

Eisner (1985) observes that ‘there is absolutely no way in which the artist can prevent the reading of the last panel before the first.’ This notion highlights a fundamental characteristic of the visual narrative form, that it is non-linear and has instantaneous accessibility. Unlike text-based narratives, where readers progress linearly from start to end, in comics, the entire page is often viewed by readers at once. Eisner further explains that an artist can only guide their readers’ eyes through the panel arrangement, size and content. Readers often tend to glance at the last panel unconditionally, skipping the first panel. It can also be a potential ‘spoiler’, but the visual openness is a unique challenge for both the comic artists and the readers who want to keep up with the suspense.

According to Cohn (2019), the building of a cognitive structure might be characterised simplistically by a three-part process: access, prediction, and updating. These involve how a reader navigates through a page layout. First, the reader accesses the panel, which happens to be a part of a larger assemblage structure that reflects the cognitive structure of how a given page is organised and how it is navigated. After this, the reader makes predictions about where to go next while reading it sequentially and encodes the meaning out of it in their long-term memories. These predictions will then allow a cycle to start, but a perceived discontinuity with the expectations will trigger an updating process. This keeps on going with each subsequent panel.

2.5 Research Gap: The Unexplored Dimensions

To highlight the research gaps through the objectives that are already discussed in Chapter 1 (Section 1.3), it is needed to emphasise what is missing or unexplored in the existing literature that is discussed above. This dissertation explores the complex cross-cultural themes embedded in visual narratives by analysing two iconic works—*Hada Bhoda* from West Bengal and *Doraemon* from Japan—alongside traditional scroll paintings such as the Bengali *Patua* and Japanese *Emaki*, representing their respective cultural contexts. While

previous studies have focused on cross-cultural perceptions, there has been limited research on how these representations in visual media have linguistic and communicative properties through their grammar and functions, and how they have evolved over time within different cultures. Additionally, the interaction between the diverse social norms and global cultural exchanges through the Bengali and Japanese cultures, respectively, has remained underexplored, especially when considering the temporal variation in the way these narratives communicate cultural themes. Moreover, the speech acts embedded within visual narratives have also remained underexplored from a communicative perspective, particularly in terms of how they convey intentions, emotions, and social dynamics non-verbally.

Cross-cultural communication allows people to avoid miscommunication and misinterpretation. It allows for opening up the possibilities of fruitful relationships. Visual communication is the most ancient ‘art culture’, starting from cave painting days (Cohn & Magliano, 2019). To make it more popular, humans have created visual narratives or picture stories that are more understandable for mass communication.

As already mentioned in Section 1.3 of Chapter 1, three specific objectives are set to explore and analyse cross-cultural perceptions through the selected visual narratives from the two distinct cultures of West Bengal and Japan. The choice of these two specific cultures stems from my Bengali origin, allowing me to compare and analyse, and Japan’s rich cultural legacy in art and manga. The Bengali comics, *Hada Bhoda*, and the Japanese manga, *Doraemon*, were selected for their widespread popularity with rich visual as well as narrative contents, distinct characteristic features, and cross-cultural appeal with societal influences.

Hada Bhoda or often known as *Handa Bhonda* (Bengali: হাঁদা-ভোঁদা) is a Bengali comic strip that was first published in 1962 and later compiled into a comic book in 2011. This comic book originated in West Bengal, India, and was created by Narayan Debnath. The comics was originally serialised in the monthly children’s magazine, ‘Shuktara’ as of September

2006. It was also made into a television series that aired for a brief time. Debnath was himself very fascinated with visual arts; coming from a family of gold retailers, art was always a part of his life in the form of designing patterns for gold jewelry (Shome-Ray, 2013). He entered the world of comics in 1962 with *Hada Bhoda*, a time when Bengali culture was still not in the world of visual narratives. *Hada* and *Bhoda* are two young boys in the comic series who reside with their fraternal aunt and uncle. *Hada* happens to be more mischievous and causes problems for others, particularly his brother, *Bhoda*. In most series, *Hada* is found to be punished for his misbehaviours.

On the other hand, *Doraemon* (Japanese: ドラえもん) is a Japanese manga series created by Fujiko. F. Fujio, which was first serialised in 1969. Similarly to Debnath (1962), Fujio is also considered to be one of the greatest manga artists in Japanese history. This beloved manga character, *Doraemon*, was created shortly after the Second World War in Japan (IGN staff, 2012). It later became even more popular in different parts of Asia with the rise of television anime series. *Doraemon* is a raccoon dog-looking cat robot from the future who can talk with humans and possesses the magical ability to pull anything fictional or nonfictional from a pouch in his belly called the miracle pocket. *Doraemon* is also considered the first anime cultural ambassador in Japan (McCurry, 2008). This robot cat came from the future to help a boy named *Nobita*. *Doraemon* uses different gadgets to solve problems for *Nobita* and sometimes his friends (*Shizuka*, *Suneo*, *Gian* and others), who are all equally mischievous like *Hada*, *Bhoda*, and their friends.

Having been created as comics and manga, respectively, *Hada Bhoda* and *Doraemon* are both products of the 20th century. These two visual narratives have brought out the two societies in front of us through their visual cultures. Despite certain similarities, these two visual narratives have significant variations, which will be revealed in the chapters of the findings. Their specific properties make them useful as stimuli for this research, along with

the Bengali *Patua* scroll paintings and the Japanese *Emaki* scroll paintings for temporal variations.

While visual narratives and animated series are widely studied and are quite popular in terms of cultural representation, there is a lack of comprehensive research that tracks Bengali and Japanese points of view from a cross-cultural and linguistic perspective. Specifically, little attention has been provided to how these representations have significance in grammatical and functional rules, exploration of both the static and dynamic elements through synchronic and diachronic data, and reflection of communicative features along with gestures and facial expressions. This research aims to fill this gap primarily with the help of a content analysis method and a qualitative approach to study cross-linguistic perceptions through visual narratives of *Hada Bhoda*, *Doraemon*, *Patua*, and *Emaki* scroll paintings and explore the temporal variation in the communication of visual culture.

This research is noteworthy as it exposes how the two cultures and their social norms are represented and how they have evolved within the visual narratives, providing new insights into global culture exchanges and the impact of visual media on societal perceptions over time.

Chapter 3

Research Methodology

3.1 Research Philosophy

This study is underpinned by a pragmatic research philosophy that prioritises practical inquiry over abstract theoretical deliberation. Pragmatism does not adhere strictly to a single ontological or epistemological stance; it allows for methodological flexibility depending on the nature of the research problem. In the context of this study, the objective is not to seek an ultimate truth or absolute reality within visual narratives but to understand how meaning is constructed and communicated through culturally embedded visual forms.

Pragmatism supports a pluralistic approach, recognising that reality can be interpreted in multiple ways, especially within diverse cultural settings. This aligns with this study's intent to explore how illustrations in different narrative traditions like the Bengali comics, *Hada Bhoda*, and the Japanese manga, *Doraemon*, serve as vehicles of meaning shaped by their respective socio-cultural milieus. Rather than seeking universal generalisations, this research focuses on contextual understanding and interpretive richness derived from the analysis of visual storytelling techniques.

Pragmatic theories of truth shift the focus away from figuring out *what* makes a statement true and instead look at how people use or understand a statement when they call it true. While these theories share some similarities with deflationary views, particularly the belief that truth isn't some deep or mysterious quality, they also go further. Pragmatists see truth as more than just a convenient way to summarise or generalise. They highlight the practical and action-oriented nature of how we talk about truth, showing that such talk plays a key role in shaping different forms of communication. From this perspective, understanding truth means paying attention to how it functions in everyday practices and discourse (Capps, 2019).

By embracing a pragmatic worldview, this research acknowledges that knowledge is situated and often constructed through interaction with context and purpose. This perspective is especially relevant in examining visual texts, where meanings are not fixed but evolve based on cultural expectations, reader interpretations, and the interplay of visual and textual elements. Pragmatism also allows the researcher to draw from both interpretivist and constructivist traditions, viewing visual narratives as cultural constructs that offer insights into how different communities perceive and represent themselves.

Although it has faced substantial criticism since its early development, the pragmatic theory of truth has persisted and even gained momentum at various points over the past century. One reason it can be challenging to evaluate this theory as a whole is that it exists in multiple distinct forms, which sometimes differ greatly from one another. Further complicating the matter, not all philosophers labeled as pragmatists have endorsed a pragmatic view of truth (Brandom, 2011), while some non-pragmatists have supported similar ideas (Dummett, 1959; Wright, 1992). On a more optimistic note, the theory has undergone significant refinement and development, making it more coherent and potentially more convincing over time. In retrospect, it becomes clear that pragmatic theories have consistently emphasised the functional role that the concept of truth serves in our practices.

The primary choice of content analysis as a methodological tool is consistent with this philosophical stance, as it provides the means to systematically explore and interpret visual material concerning broader societal themes. The philosophy guiding this research, therefore, is not preoccupied with abstract philosophical binaries but is firmly anchored in the practical objective of understanding cultural expression through visual media. This allowed the study to remain open, adaptive, and sensitive to the nuances that emerge from cross-linguistic and cross-cultural comparison.

In terms of its axiological stance, this study adopts a value-aware but reflexively neutral position. While it does not claim complete detachment from the cultural materials under investigation, it strives to be critically self-aware of the researcher's influence during interpretation. Pragmatism, as a philosophical orientation, does not insist on absolute objectivity but encourages an awareness of values and their impact on the research process. Rather than assuming a 'value-free' stance, an ideal often critiqued in qualitative inquiry, this research maintains a value-neutral position that recognises the interplay between researcher subjectivity and interpretive validity. This is particularly relevant given the cross-cultural nature of the study, where recognising one's own cultural lens is essential in avoiding ethnocentric biases.

To ensure a balanced and credible analysis in a subjectively interpretive framework, triangulation was incorporated through methodological complementarity. While the core data were analysed using qualitative content analysis, a small component of ethnographically informed fieldwork, including semi-structured interviews and observational engagement with *Patua* artists during a field visit to Pingla, offered valuable socio-cultural context. These field insights supported the content analysis by grounding visual interpretation in lived experience and cultural practice, thereby contributing to a richer, more nuanced understanding of meaning-making across cultures. This integration of limited fieldwork also aligns with the pragmatic philosophy, allowing methodological decisions to be driven by the research questions and contextual demands rather than theoretical purity.

Therefore, the research philosophy adopted here values empirical grounding, contextual understanding, and methodological openness—hallmarks of pragmatic inquiry that support the study's overarching goal of decoding culturally specific modes of visual communication.

3.2 Research Approach

This study adopted a qualitative research approach to explore and analyse the cross-cultural differences and similarities in visual narratives from two distinct cultural contexts, Bengali and Japanese. The analysis concerns topics like grammar, structure, and social and temporal contexts, along with the different communicative features that are ingrained within the illustrations of *Hada Bhoda*, *Doraemon*, *Patua*, and *Emaki* scroll paintings. These visual narratives are analysed not only in terms of their grammatical, structural and artistic expressions but also as cultural artifacts that reflect and shape societal values such as family, education, food, and other social surroundings.

Pragmatism as a research paradigm avoids engagement with complex metaphysical ideas like truth and reality. Rather, it embraces the idea that there may be one or multiple realities that can be explored through empirical investigation (Creswell & Clark, 2011). Primarily, a content analysis methodology was used because, as a pragmatic researcher, I am concerned with solving real-world problems and creating practical knowledge. This method is consistent with my desire to understand the context in which the research is conducted. Using content analysis methodology, I carefully studied the visual narratives of *Hada Bhoda*, *Doraemon*, *Patua* and *Emaki* scroll paintings, respectively, paying great attention to the social, cultural and organisational elements that influence my research problem and its solutions. This method allowed me to extract important patterns and insights from the selected visual narratives, resulting in a more complete knowledge of how visual culture is communicated and how it reflects broader cultural values and conventions.

Moreover, to analyse visual material, a semi-structured interview and a brief field observation were conducted with *Patua* artists (especially Ranjit Chitrakar) during my field visit to *Naya, Pingla*. This element of the study, though limited in scope, was essential to support the second objective, exploring the lived and performative dimensions of traditional visual storytelling. Incorporating this ethnographically informed method allowed the study to

capture aspects of meaning that might not be accessible through content analysis alone, thereby enhancing the interpretive depth and ensuring cultural sensitivity in the analysis. The integration of multiple qualitative tools reflects a pragmatic approach, methodologically flexible and grounded in the research objectives, focusing on what works best to explore the problem in its real-world complexity.

Because we believe pragmatic matters take precedence over philosophical ones, we do not overburden the reader with discussions of epistemological difficulties. Epistemology is the study of the philosophical foundations that underpin distinct approaches to knowledge formation (Bryman, 2004). Epistemological analysis is useful for research methodologists to understand the implications of using specific methodologies. As a pragmatic researcher, the approach of this study focused on addressing real-world problems and producing actionable knowledge by primarily adopting a content analysis methodology. The selected visual narratives, as mentioned earlier, are systematically analysed in this study. Content analysis provided a flexible yet robust framework, ensuring that the findings are contextually relevant and practically applicable.

By analysing and exploring visual narratives through a theoretical framework that will be discussed thoroughly in this chapter, this research aims to elucidate how visual and linguistic elements combine to construct compelling narratives, utilising grammatical and functional roles and conveying temporal variations, communicative features, emotions, and expressions, with the help of cultural nuances. By focusing on these aspects, this research was able to explore the nuances of cultural representation and narrative expressions without imposing any external influence, contributing to a deeper comprehension of cross-cultural visual communication.

The purpose of the study lies potentially in the exploration of cross-cultural notions present in the illustrations through their grammar, structure, function, temporal contexts and

communicative elements of the selected visual narratives that are already discussed in the objectives of the study. Altogether, it can be said that visual narratives can act both as a mirror and a looking glass into society. It can therefore serve as a dynamic medium that could also be used as a universally applicable mode for various academic purposes in the long run. If things and events had any intrinsic true essence, there would have been no space for cross-cultural variations: this exploration highlights that visual narratives reflect culturally specific interpretations rather than truths, and it is discussed throughout this dissertation and analysed well with the help of a rich variety of findings.

Using a theoretical framework, this research mainly applies a content analysis method along with a short interview part to explore and look into the culturally diverse visual narratives of *Hada Bhoda* and *Doraemon*, as well as *Patua* and *Emaki* scroll paintings. This chapter describes the research design and methodology, encompassing the data collection process and data analysis procedures necessary to achieve the objectives of this study.

3.3 Theoretical Framework

In the domain of visual narratives, Bengali and Japanese visual narratives provided this research with an extensive mosaic of narrative styles and social commentary. Understanding the structural and functional aspects, the temporal aspects and the speech acts with multimodality revealed how cultural values and social settings are woven into visual narratives. As already mentioned earlier in Chapter 1, the general objective of my dissertation is as follows:

To study cross-cultural perceptions through selective visual narratives and explore if and how the temporal variation in the communication of visual culture varies over time.

The theoretical framework for this study draws on concepts from semiotics and Visual Linguistics to analyse cross-cultural visual storytelling. Using a primarily content analysis method, within a qualitative approach, this framework relies on thematic coding to

categorically and systematically analyse the specific objectives under the general objective and elicit answers to the research. To analyse the cross-cultural notions of the visual elements present within *Hada Bhoda*, *Doraemon* as well as *Patua* and *Emaki* scroll paintings, a cross-linguistic analysis is applied with the help of a theoretical framework to bring out the Bengali and the Japanese linguistic values from the visual narratives, respectively. This framework integrates Chomsky's structural theory (1957), Halliday's functional theory (1916), Saussure's diachronic and synchronic analyses (1916), and Austin's speech act theory to explore the linguistic features in visual narratives. By applying these linguistic theories, the thematic coding identifies the patterns in the narrative organisations, cultural symbolisms, temporal variations, and communicative elements that are unique to each cultural context.

For the first objective, the visual narratives, *Hada Bhoda* and *Doraemon*, are cross-culturally analysed using grammar, structure and social settings to explore how visual culture functions.

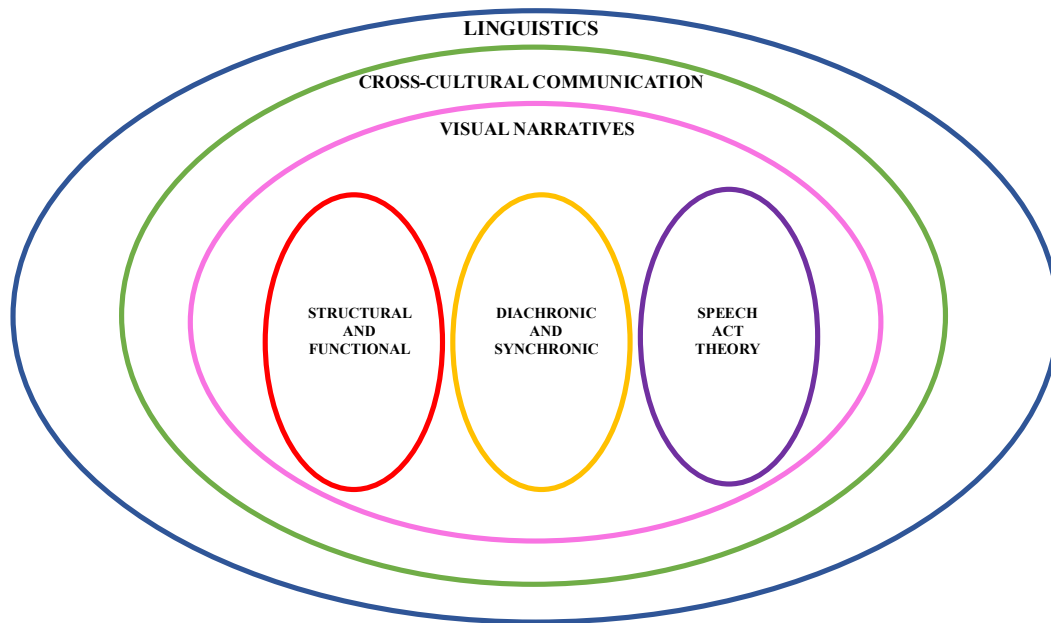
For the second objective, the study looked into how the frameworks of *Patua* and *Emaki* scroll paintings differ from the visual narrative grammar of Bengali comics and Japanese manga, respectively.

And lastly, for the third objective, the study explored how visual culture is communicated in the visual narratives of *Hada Bhoda* and *Doraemon* through facial expressions, ideophones and prominent drawing styles of the two cultures.

Under this viewpoint, the theoretical framework emphasises the depths and complexities of Visual Linguistics within visual narratives of the two selected cultures by incorporating a few linguistic ideas that have already been mentioned briefly in Section 1.3.4 of Chapter 1. By exploring visual narratives through these theoretical lenses, we can acquire a better knowledge of how visual and linguistic aspects work in the same way to generate captivating plots, which enhances our comprehension of cross-cultural communication and visual culture.

This approach can help strengthen our insights into visual storytelling by revealing the subtle ways in which culture shapes narrative expressions:

Figure 4. Theoretical Framework



Source: Developed by the author (Chattopadhyay, 2023)

By incorporating these interdisciplinary perspectives of Linguistics, this framework not only enriches the study of visual narratives but also sheds significant light on broader concerns in visual communication or Visual Linguistics. It provides a broad understanding of how visual and textual elements work together with a bimodal approach to generate unique and culturally relevant stories. This method highlights the subtle ways in which visual narratives function as a medium of human expression, emphasising their importance in communicating complex values and social conventions.

Cross-cultural communication is a significant factor in Linguistics, focusing on how people from different cultural backgrounds use language and non-verbal cues to interact and understand one another. Moreover, this paradigm or framework emphasises the significance of visual narratives in cross-cultural communication, illustrating how they bridge cultural divisions and promote mutual comprehension. We may appreciate the richness and

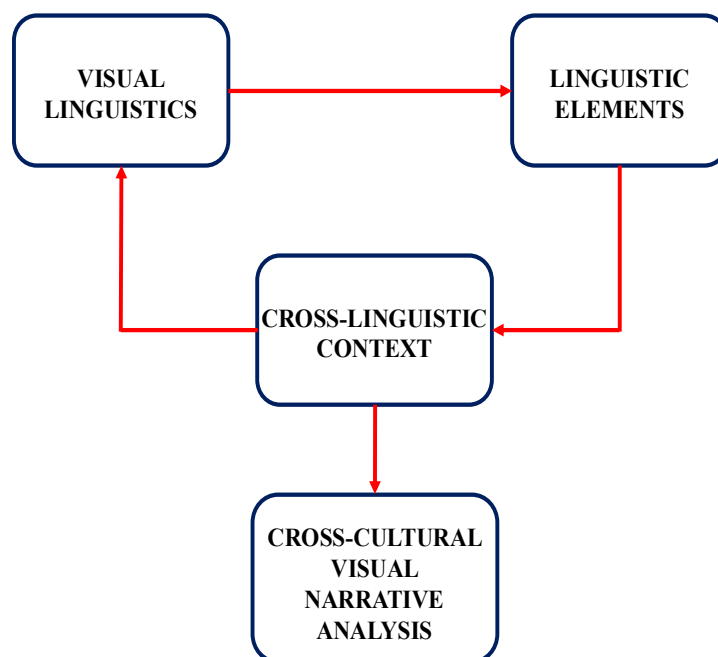
complexities of visual storytelling by looking at the grammar, structure, functional roles, temporal aspects, and communicative features included within it; this heightened interest not only affirms the importance of visual storytelling in cultural discourse but also encourages a greater understanding of the various ways in which different cultures express their distinct experiences and perspectives through visual art.

This interdisciplinary approach of cross-linguistic analysis emphasises the value of visual storytelling in representing and preserving cultural identities. Therefore, this framework supports the general objective along with the specific objectives of the study by demonstrating how this thematic coding can uncover the cultural resonances in visual narratives, highlighting the role of visual culture in cross-cultural communication.

3.4 Conceptual Design

Based on the theoretical framework in Section 3.3, the conceptual design visually organises the operational elements that guide my analysis of *Hada Bhoda* and *Doraemon* within a cross-cultural and cross-linguistic lens. Where the theoretical framework draws from linguistic theory to explain the why and what, the conceptual framework illustrates the how, as in how these linguistic theories are practically applied to decode, categorise, and analyse meaning across visual narratives:

Figure 5. Conceptual Design



Source: Developed by the author (Chattopadhyay, 2023)

This conceptual framework in Figure 5., translates the foundational linguistic theories into the interconnected analytical components that operationalise the research objectives. Central to the framework is the Cross-Linguistic Context, which mediates the dynamic relationship between Visual Linguistics and Linguistic Elements. The interaction between these domains reflects the application of theoretical insights—such as Chomsky’s structuralism, Halliday’s functions, Saussure’s temporal and comparative perspectives, and Austin’s speech acts—to visual storytelling across cultures.

Visual Linguistics refers to the non-verbal elements, such as panel structures, facial expressions, gestures, social contexts, and artistic styles that are present in both *Hada Bhoda* and *Doraemon*. These are not merely illustrations but culturally embedded communicative units that function like language. Visual linguistic elements are crucial in shaping how narratives are perceived and interpreted, especially across different cultural frameworks.

Linguistic Elements encapsulate the textual components embedded in the visual narratives— dialogues, ideophones, onomatopoeia, captions, and written expressions. These elements reflect syntactic structures, pragmatic uses, and culturally specific semantics that shape narrative understanding and emotional resonance.

The Cross-Linguistic Context serves as the interpretive space where Bengali and Japanese linguistic and visual elements are systematically compared. It allows for the recognition of both convergence and divergence in narrative structures, ideational content, and expressive strategies shaped by language and culture.

The culmination of these interconnected parts leads to the Cross-Cultural Visual Narrative Analysis. Here, insights from the interactions of linguistic and visual systems are categorised through thematic coding, informed by the theoretical foundations. This step supports the broader aim of uncovering how cultural values, temporal concepts, and communicative norms are visually and linguistically narrated in distinct yet relatable ways in *Hada Bhoda*, *Doraemon*, *Patua*, and *Emaki* scroll paintings.

This framework thus ensures that the theoretical foundation is applied meaningfully to interpret the visual and linguistic complexities embedded in cross-cultural storytelling through the medium of visual narratives, enabling a rigorous and nuanced comparative analysis.

3.5 Research Method

A content analysis methodology was primarily used to accomplish the goal of this study. As already mentioned in Chapter 2, Neuendorf (2017) defines content analysis as an objective, systematic, and quantitative approach to examining the features of a message. Whereas, Krippendorff (2004) stated that reading any text inherently involves a qualitative approach, even if certain aspects are eventually quantified. Fraenkel and Wallen (2006), on the other

hand, described content analysis as a research method through which scholars study human behaviour indirectly by analysing the messages and communications that people produce.

The primary goal of content analysis is to study social behaviours and experiences without influencing them. This could be accomplished by examining any form of visual or written human communication, such as items appearing in television commercials, novels, newspapers, magazines, television, speeches and many others (Fraenkel and Wallen, 2006). This study employed a qualitative content analysis. The qualitative factor allowed this research to hear the voices of the readers of the two selected visual narratives through the artist's or the writer's perceptions. By using the content analysis method, this research was able to study the behaviours and experiences of the two different cultures without any external influences. In other words, the selected visual narratives (*Hada Bhoda*, *Doraemon* and *Patua* and *Emaki* scroll paintings) were evaluated and analysed by focusing on the three objectives that have already been mentioned earlier (see, Section 1.3, Chapter 1). A qualitative content analysis methodology was applied to the following aspects:

For the first objective, which involves cross-cultural analysis of the visual narratives of *Hada Bhoda* and *Doraemon* through the grammar, structure, and social contexts of Visual Linguistics, the structural and functional theories of Linguistics by Chomsky (1957) and Halliday (1961) were used, respectively. The research first focused on the ordering of the panels in both the visual narratives and tried to bring out the cultural importance through it. The ordered panels gave an outlook on the grammar of visual narratives—the very basics of how to read a Bengali comics or a Japanese manga. The grammar of visual narratives helps shape and convey the narrative structure of the stories. The set of rules or the organisation of framework created by the rules governs the narrative structures through illustrations and the choice of sentences. Therefore, the study looked into the differences in illustrations, and as a result of it, the similarities and differences in the narrative structures of *Hada Bhoda* and *Doraemon* were found, which eventually brought out the variations in the visual narratives

through cross-cultural specifications. This gave rise to variations in the social contexts of the visual narratives. With the help of the structural and functional theories of Chomsky (1957) and Halliday (1961), the study analysed and explored how bimodal communication in visual narratives is used not only to convey information grammatically but also to form storylines functionally. The functions of these visual narratives are analysed through cultural reflections on family dynamics, friendly exchanges, school conversations, sportsmanship and culinary culture.

For the second objective, the research focused on the temporal aspects in the communication of visual culture over time. For this, scroll paintings of both cultures (*Patua* scroll paintings of West Bengal and *Emaki* scroll paintings of Japan) were cross-analysed with the help of Saussure's (1916) diachronic and synchronic approaches. To achieve the goal of this objective, I visited *Naya, Pingla*, a popular village in West Bengal with a strong legacy of *Patua* scroll paintings, and gathered information. I spoke to the local artists, saw their procedures, and collected important data for the study. For *Emaki* scroll paintings, a few relevant data points were found on the internet. The study looked into the iconicity of the achronological storytelling of differently ordered panels, which are present in both *Patua* and *Emaki* scroll paintings, respectively. It also found out how the natural gutters in these scroll paintings are different from what they are in the modern visual narratives of Bengali comics, *Hada Bhoda*, and the Japanese manga, *Doraemon*, respectively.

To supplement this textual analysis with ethnographic insights, additional data collection was undertaken during the field visit:

While qualitative content analysis was the primary methodology, this study also incorporated a semi-structured interview and field observations to support this research objective. Specifically, I conducted a brief interview with Ranjit Chitrakar and his family (*Patua* artists) during my field visit to *Naya, Pingla* in West Bengal, India. These interviews were informal

and narrative in nature, allowing the artists to share insights about their storytelling traditions, drawing techniques, and the transmission of knowledge across generations. Although this component was relatively small in scale, it provided valuable contextual depth to the analysis of *Patua* scroll paintings.

Additionally, the observational component complemented the content analysis by grounding the visual texts in their real-world cultural practices, making the methodology more holistic and ethnographically informed. By combining textual analysis with field interviews and observation, the study adopted a multi-method qualitative approach. While the core analytical method remained content analysis, the integration of interview and observational data enabled a richer understanding of the socio-cultural contexts in which these visual narratives were created and consumed.

For the third objective, the research explored thoroughly how visual culture is communicated with the help of the visual narratives of *Hada Bhoda* and *Doraemon* through the different communicative acts, facial expressions, different ideophones or onomatopoeic words, and the different drawing styles with the help of Austin's (1959) speech act theory. The study carefully observed and analysed the different facial expressions found in both *Hada Bhoda* and *Doraemon* to better grasp how emotions and cultural nuances are expressed visually. By exploring a variety of emotions, from happiness and surprise to anger and sadness, this study was able to identify the intricacies of how each culture visually communicates emotional states. It also looked into the different culture-specific ideophones and how some of them differ even when the pictures are similar. Furthermore, it also explored the prominent drawing styles through these observations.

3.6 Sampling Design

In this research, the sampling design is based on purposive sampling, which is a non-probability sampling technique that has allowed the deliberate selection of the Bengali and

Japanese visual narratives and scroll paintings, respectively, based on their relevance to the research objectives. The benefits of purposive sampling are many, such as targeting the population of particular interest or making the sample more uniform, reducing the variability between subjects, and increasing the likelihood of finding statistical significance. It can also exclude participants at risk of severe adverse events (Andrade, 2020). The selected samples used in this research, *Hada Bhoda* from West Bengal (India) and *Doraemon* from Japan, respectively, were chosen because of their significant cultural impacts, widespread popularity, and rich visual content, giving them ideal outlooks for cross-cultural analysis. Along with the visual narratives, *Patua* scroll paintings from West Bengal and *Emaki* scroll paintings from Japan were included as historical stimuli to explore how visual narrative structures have evolved over time. These scroll paintings were selected for their symbolic and representational roles in the respective visual cultures of Bengal and Japan, tracing the temporal variations and developments in picture stories.

Additionally, the selection of these two cultural traditions was shaped by both personal familiarity and academic support. As a researcher of Bengali origin, I was naturally positioned to access, interpret, and critically engage with Bengali visual narratives such as *Hada Bhoda* and *Patua* scrolls. The Japanese materials, including *Doraemon* and *Emaki* scrolls, were explored with guidance and support from my co-supervisor, whose expertise helped facilitate deeper understanding and access.

Furthermore, the Bengali and Japanese traditions share certain parallels in many factors, while also differing significantly in structure, artistic conventions, and cultural symbolism. These shared and divergent elements enhanced the relevance of this cross-cultural comparison and provided a strong foundation for a nuanced cross-linguistic analysis.

The universe for this study includes the entire body of Bengali and Japanese visual narratives, particularly comic books and traditional scroll paintings that have evolved from their

respective cultural storytelling traditions. This encompasses a wide range of visual storytelling forms that reflect the socio-cultural, linguistic, and aesthetic sensibilities unique to each culture. In the Bengali context, this includes serialised comic books, picture stories, and hand-painted *Patua* scrolls, which have historically served both entertainment and pedagogical purposes. Similarly, in the Japanese context, the universe comprises various genres of manga as well as *Emaki* scrolls, which represent one of the earliest forms of Japanese illustrated storytelling. These diverse formats, spanning oral traditions, print culture, and visual literacy, constitute the broader population from which the purposive samples for this research have been drawn. This includes:

- Bengali comic series from the 20th century onward, with particular focus on *Narayan Debnath's* works such as *Hada Bhoda*,
- Japanese manga series such as *Fujiko F. Fujio's Doraemon*,
- Traditional *Patua* scroll paintings of Bengal, and
- Historical *Emaki* scroll paintings of Japan.

This universe spans multiple genres and periods, but for a focused cross-cultural analysis, a purposive sample was drawn from within this broad range based on specific selection criteria.

From this broader universe, a purposive sample was drawn using clear criteria based on cultural impact, accessibility, visual structure, and relevance to the research questions. The selected sample includes:

1. *Hada Bhoda* (Bengali Comic Series):

Source: *Hada Bhoda Samagra* (Debnath, 1962, collected volume)

Sample Size: The entire collected volume was used for analysis.

Figures Used: Figures 14–16, 20–22, 25–27, 31–32, 35–37, 54, 56, 58, 63–64, 71, 73–77, 83–86 (and others).

Rationale: *Hada Bhoda* is one of the most iconic Bengali comics. Its humour, visual grammar, and cultural expressions make it an ideal case for cross-cultural comparison.

2. Doraemon (Japanese Manga Series):

Source: *Doraemon* Volumes 1–5 and *A Selection of Touching Stories* (Fujio, 1974)

Sample Size: 6 volumes in total

Figures Used: Figures 17–19, 23–24, 28–30, 33–34, 38–40, 60, 62, 69–70, 72, 78–82, 83–86, and more.

Rationale: *Doraemon* is one of the most globally recognised Japanese manga, offering rich visual grammar, linguistic features, and culturally embedded narratives relevant to children and adults alike.

3. Patua Scroll Paintings (West Bengal, India):

Source: Primary data collected during fieldwork in *Naya, Pingla*, West Bengal.

Artists: Scrolls were primarily created by *Ranjit Chitrakar*, a well-known *Patua* artist. Additional scrolls were made by his father-in-law and a few other local artists, whose names were not recorded. Brief conversations were also held with his son and wife, who are part of the scroll-making tradition.

Sample Size: 7 scroll paintings were analysed.

Figures Used: Figures 42–48, 53, 55, 57

Rationale: The selected scrolls reflect a variety of narrative techniques and panel structures, including chronological and achronological ordering, single and multiple panels, and different reading paths. These scrolls serve as key cultural artifacts for understanding the evolution and grammar of Bengali picture storytelling.

4. Emaki Scroll Paintings (Japan):

Source: Secondary data obtained from digitised archives of museum collections (public domain and scholarly sources).

Sample Size: 5 scroll paintings

Figures Used: Figures 49–52, 59, 61

Rationale: *Emaki* scrolls like *Genji Monogatari Emaki* and *Chōjū-giga* demonstrate early Japanese approaches to visual storytelling. Their format, narrative progression, and integration of text and imagery align with the analytical framework used in this research.

While the universe of visual narratives analysed in this study includes entire volumes of *Hada Bhoda* and *Doraemon* and several scroll paintings, not all picture materials were analysed in equal depth. The figures presented in this study were purposively selected based on their relevance to the research objectives and theoretical framework.

Specifically, panels and scroll segments were chosen based on their:

- Exemplification of narrative transitions {e.g. moment-to-moment, action-to-action, scene-to-scene, etc. as per McCloud's (1993) categories},
- Representation of unique visual-linguistic features, including panel structure, reading direction, facial expressions, and speech/text integration,
- Illustration of culturally embedded themes, and
- Contrastive value, allowing for cross-cultural comparison between Bengali and Japanese visual storytelling modes.

Additionally, *Patua* and *Emaki* scrolls were selected to reflect diverse structural techniques such as chronological sequencing, the use of gutters, and the integration of visual and textual elements. Scrolls with complex panel paths and those covering culturally significant narratives were prioritised for inclusion.

Therefore, the selection of specific images (figures) was not random but made from the researcher's point of view, guided by the theoretical framework of this study, which is a systematic Visual Linguistic analysis. This ensured that the chosen visual data supported a

focused and coherent comparative study cross-culturally.

Under this condition, the criteria for selecting these specific visual narrative samples from the two cultures are as follows:

- i. Visual Grammar and Linguistic Features:** Both cultures exhibit these selected visual narratives with distinct visual structure, grammar, and functions, including linguistic features revealing temporal variations and communicative acts. These grammatical features reflect unique cultural conventions and storytelling techniques that shape these narratives into communicable and comprehensible contexts in Bengali and Japanese picture stories, respectively.
- ii. Cross-cultural Relevance:** The cross-cultural relevance of this research lies in the analysis of these chosen visual narratives and how *Hada Bhoda* and *Doraemon* have justified their cultures, respectively, in reflecting and communicating cultural values and social norms accurately. Additionally, *Patua* and *Emaki* scroll paintings have also contributed equally in analysing and comparing the two cultures, highlighting how picture stories serve as a medium for cross-cultural exchange, making their way into a deeper comprehension of both shared and divergent cultural perceptions and beliefs.
- iii. Widespread Popularity and Prominence:** The visual narratives, *Hada Bhoda* and *Doraemon* are both culturally popular and recognised; the latter is globally accessible as well. Therefore, it ensures a broader representation of visual culture in both societies and beyond by generalising the results to a larger aspect of the cross-cultural analysis of visual narratives.
- iv. Temporal Variation:** The evolution of the visual narratives of the two respective cultures, Bengal and Japan, over time, includes the narratives of *Hada Bhoda* and *Doraemon* along with the ancient Bengali *Patua* scroll paintings and Japanese *Emaki* scroll paintings, providing both synchronic and diachronic perspectives. This analysis has

been facilitated by the fact that both Bengali and Japanese cultures share a tradition of scroll paintings as ancient narrative forms capturing the cultural richness in both societies.

Therefore, the purposive sampling design applied in this study ensures a focused and culturally relevant analysis, even though it may have limitations in the findings, which are beyond the selected visual narratives and cross-cultural contexts.

3.7 Method of Data Collection

The data collection in this dissertation primarily follows a qualitative content analysis approach that focuses on the selected visual narratives from West Bengal (India) and Japan. In addition, a short field visit was conducted for *Patua* scrolls to support the analysis with contextual insights. As mentioned earlier, the primary sources of data or the stimuli of this study include the Bengali comics *Hada Bhoda* and the Japanese manga *Doraemon*, followed by *Patua* scroll paintings from West Bengal and *Emaki* scroll paintings from Japan, for comparing and analysing the temporal aspects. All of these were selected for their cultural significance and rich visual content.

The qualitative approach in this study enabled an in-depth analysis and exploration of the linguistic features with cultural contexts for these visual narratives. The visual narratives were carefully explored and analysed to identify patterns in panel organisation, social contexts, temporal aspects, communicative acts, facial expressions, and other communicative features, which provided a cultural lens for understanding cross-linguistic perceptions and how visual storytelling methods have evolved across the two cultures.

The data was collected through textual and visual content analysis, comparative analysis, and thematic coding, by dividing them according to the objectives of the study, respectively:

- i. Textual and Visual Content Analysis:** The visual narratives were broken down into their respective components and were analysed through their panels, gutters, and story

sequences in order to understand their grammatical, structural, functional, and other linguistic features.

- ii. **Comparative Analysis:** A cross-cultural comparison was carried out between the selected Bengali and Japanese visual narratives, with a focus on how these picture stories depict different linguistic and social elements.
- iii. **Thematic Coding:** The collected data were organised into different linguistic themes as per the theoretical framework, such as aspects related to grammar, structure and functions, diachronic and synchronic variations, and speech act theory. These themes were then compared across the two cultures through the visual narratives.

3.8 Method of Data Analysis and Interpretation

The method of data collection provided a comprehensive understanding of the study of how the visual narratives from the two different cultures can communicate their social and cultural contexts by analysing them with the help of linguistic features. The data analysis for this study was conducted with the help of the qualitative content analysis method, focusing on the cultural, linguistic, and thematic elements of the selected visual narratives, as already mentioned earlier. This study aims to uncover the patterns and differences in how these visual narratives (*Hada Bhoda*, *Doraemon*, *Patua* and *Emaki* scroll paintings) reveal their linguistic properties and reflect cultural norms, respectively.

The following steps included the data analysis process for this study:

- i. **Data Familiarisation:** The selected visual narratives were reviewed closely, and the textual and visual contents were observed in detail to develop a clear comprehension of their underlying linguistic and cross-cultural properties.
- ii. **Thematic Coding:** The key themes, according to the objectives and the theoretical framework of this study, such as grammar, structure, and functions, temporal aspects, and speech act theory, were identified through coding. This thematic coding helped categorise and organise the data into meaningful groups as per the objectives of this

study, highlighting significant cultural differences and similarities among the chosen visual narratives of West Bengal and Japan. This coding involved mainly observation and a short field visit for *Patua* scroll paintings for the second objective, which deals with temporal aspects. These *Patua* scroll paintings were compared and analysed through the iconicity and semantic differences to those of the *Emaki* scroll paintings that were found on the internet.

iii. Cross-cultural Comparative Analysis: The theoretical framework mentioned in Section 3.2 of this chapter was applied to analyse the visual narratives in terms of their cultural relevance to explore the narrative structures through their grammatical and functional aspects, the temporal aspects of visual narrative evolution from scroll paintings to Bengali comics and Japanese manga, and the speech act theory through communicative aspects, facial expressions, use of ideophones, and drawing styles. The analysis was designed to reveal both shared elements and culturally specific interpretations of linguistic values in the two cultures.

iv. Synthesis of Findings: The insights from the thematic coding held by the theoretical framework and cross-cultural comparison with the help of linguistic properties were synthesised to conclude the cross-linguistic analysis and significance of the visual narratives. This analysis emphasised how the visual narratives serve as a lens of Linguistics into the two distinct cultural perceptions of West Bengal and Japan.

In order to interpret the data meaningfully within and across cultural contexts, this study drew upon two complementary analytical perspectives from the social sciences—emic and etic. These frameworks are particularly useful in cross-cultural research where insider understanding must be balanced with comparative objectivity. The emic/etic distinction originates from anthropology and offers complementary lenses: emic to understand meanings significant within a culture, and etic to provide generalisable comparisons across cultures (Mostowlansky & Rota, 2020/2023). These dual perspectives helped balance internal cultural

meanings with comparative linguistic and narrative analysis. Therefore, to strengthen the interpretive depth of the analysis, both perspectives were employed in this study:

Emic interpretation involved interpreting certain data from within the cultural context, particularly the *Patua* scrolls. This was supported by the brief field visit to *Naya, Pingla*, where insights were gathered from the artist *Ranjit Chitrakar* and his family. Their explanations helped contextualise symbolic meanings, panel layouts, and oral narrative functions intrinsic to *Patua* storytelling.

Etic interpretation was applied across all visual data, using the theoretical framework based on linguistic theories. Analytical tools such as McCloud's (1993) transitions of comics and Cohn's (2013) visual grammar model were applied to *Doraemon*, *Hada Bhoda*, and the scroll paintings to draw cross-cultural parallels and contrasts.

While *Hada Bhoda* and *Patua* scrolls benefited from my Bengali background and cultural familiarity, which can be seen as the emic insight, *Doraemon* and the *Emaki* scrolls were primarily analysed through an etic lens, supported by secondary literature, translated manga volumes, and academic supervision.

Therefore, a nuanced comprehension was provided by the methodological approach to analyse the data of how *Hada Bhoda*, *Doraemon*, and the scroll paintings of the two cultures represent and communicate cross-linguistic values by offering a detailed cross-comparison with the help of the linguistic properties present in the theoretical framework.

3.9 Limitations and Delimitations

It is essential to rethink the global research agenda and emphasise transparency in research designs. Authors should clearly state research limitations, delimitations, and assumptions, as excluding these elements can undermine the reliability of the study. When these terms are

thoroughly discussed, studies gain higher credibility and are more likely to be referenced as benchmarks for future research (Theofanidis & Fountouki, 2018).

3.9.1 Limitations

This study provides valuable insights into the cross-linguistic analysis of visual narratives, while certain limitations are there. They are as follows:

- i. Limited Scope of Visual Narratives:** The study focuses on the two specific visual narratives, *Hada Bhoda* and *Doraemon*, along with the Bengali *Patua* and the Japanese *Emaki* scroll paintings, respectively. Although these are rich sources and representatives of the Bengali and the Japanese cultures, respectively, yet, they do not encompass the entire spectrum of visual narratives from these two specific regions. Other narratives from the same cultural backgrounds may offer different perspectives on cross-cultural representation.
- ii. Language and Cultural Nuances:** The study might have encountered certain challenges in accurately interpreting a few cultural nuances, especially those in the Japanese manga, *Doraemon*, since English translated forms were studied. These translations may or may not fully capture the fine details of the original language based on the cultural context, potentially leading to a loss of meaning or misinterpretation of certain visual elements.
- iii. Subjectivity in Interpretation:** Subjectivity in data analysis may influence the interpretation of visual and textual elements, thematic coding through the theoretical framework, and other comparative results. The analysis wholly depends on my (the researcher's) observations and interpretations of symbolic and visual cues, which can differ from person to person.
- iv. Temporal Aspects:** A part of the research focuses on the historical and contemporary significances of *Patua* and *Emaki* scroll paintings in comparison to *Hada Bhoda* and *Doraemon*, respectively. However, these narratives represent specific timelines, and their historical relevance may evolve over time, making it challenging to generalise findings to

future or past visual narratives. While a short field visit to *Naya, Pingla* (West Bengal) was made to observe the *Patua* scroll paintings, such direct observation was not applied to *Emaki* scroll paintings in Japan, which may create a slight disparity in the depth of cross-linguistic analysis.

3.9.2 Delimitations

It was an intentional act to narrow down the focus and scope of this study, ensuring manageable and meaningful results. Therefore, the delimitations are as follows:

- i. Selection of the Visual Narratives:** This research is delimited to a few visual narratives for the findings—*Hada Bhoda* and *Doraemon* mainly, and the Bengali *Patua* and the Japanese *Emaki* scroll paintings. These picture stories were purposefully selected for their cultural richness, distinct narrative structures and linguistic features. The research does not include other visual narratives to maintain a proper focus for comparison.
- ii. Cross-Linguistic Comparison:** The analysis is restricted to comparing the Bengali and Japanese cultures; This allows a focused exploration of the selected contexts. Therefore, it also broadens the representation of visual culture in both societies through linguistic aspects and extends its relevance by applying the findings more universally from a linguistic point of view.
- iii. Use of Thematic Coding with the Theoretical Framework:** The theoretical framework enabled the thematic coding approach that allowed the study for a detailed breakdown of visual linguistic features, ensuring a well-organised and thorough qualitative analysis.
- iv. Exclusively Visual Linguistic Analysis:** The study and its analysis are wholly confined to visual linguistic dimensions within the selected visual narratives. As a result, the focus was fully on visual grammar, structure, functions, temporal aspects, and communicative features, offering an in-depth analysis of Visual Linguistics through the selected picture stories. By focusing on these linguistic features, the dissertation provides valuable insights into how visual language functions as a structured form of communication across

different cultures. The study also intentionally excludes audience feedback as it focuses strictly on the visual linguistic aspects present within the selected visual narratives.

v. Contribution beyond Visual Linguistics: The study advances the field of Visual Linguistics by applying its linguistic aspects to visual media and demonstrating how visual storytelling transcends verbal language and provides a basis for future research in cross-linguistic variations in visual culture. This study offers Linguistics, Semiotics, and other visual studies an enriching analysis through the medium of meaning-making. It can also enrich our comprehension beyond Linguistics in the fields of Media, Education, and Psychology by providing significant insights into how visual storytelling can be practiced universally in multicultural contexts.

In this context, the study ensures a concentrated analysis of cross-linguistic perceptions within the selected visual narratives, allowing for a thorough exploration of the themes and objectives mentioned earlier.

Chapter 4

Findings (Objective 1): A Cross-Linguistic Analysis of Visual Narratives through the Grammar, Structure and Social Contexts

4.1 TGL and SFL in Visual Narratives

To cross-culturally analyse the visual narratives of *Hada Bhoda* and *Doraemon* through the grammar, structure, and social contexts of visual culture used in these visual narratives.

For the first objective of this study, an attempt was made to cross-culturally analyse the Bengali comics, *Hada Bhoda*, and the Japanese manga, *Doraemon*, using grammar, structure, and functions through the social settings of these two visual narratives. Chomsky's (1957) TGL and Halliday's (1961) SFL were used to determine the syntactic structures and principles that regulate panel ordering into functional and cultural significances. This objective focuses on the cultural importance of panel ordering, the differences in drawings in narrative structure, and how the social contexts are presented. The content analysis method and qualitative research provided a complete framework for studying the intricate interaction of these visual forms.

4.2 Structure and Grammar of Visual Narratives (*Hada Bhoda* and *Doraemon*)

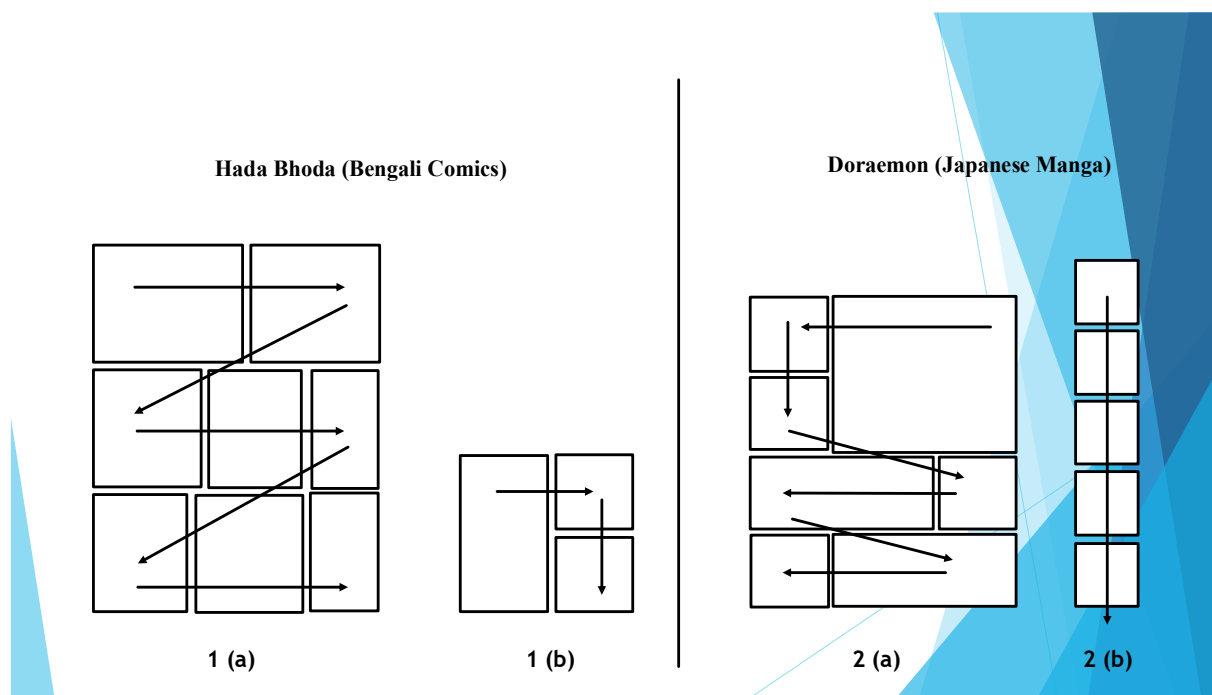
The first nine pages of McCloud's (1993) 'Understanding Comics' are totally devoted to the definition of comics as a medium. He developed the very definition by putting visual narratives as sequential art—a collection of still images placed sequentially to construct a story or to convey emotions; therefore, making it a very common technique used in visual narratives. A typical visual narrative page will consist of multiple panels arranged in a grid-like structure. This can vary in size and according to the artist's choice; this whole narrative structure helps in establishing a narrative flow and a rhythm of the whole picture story. The bimodal form of visual narratives, i.e., the visual and verbal methods used to create visual

narratives like comics and manga, forms the grammar of picture stories. Visual elements such as panels, gutters, speech bubbles, captions, and other such depictions are used together to portray a tale or narrative.

McCloud (1993) categorises the changes between panels as part of the grammar of visual narratives. These changes help the readers to shift semantically from one panel to another, thereby forming the narrative in a comprehensive way.

The following diagram illustrates the arrangements of the panels in the visual narratives of *Hada Bhoda* (Bengali Comics) and *Doraemon* (Japanese Manga), respectively:

Figure 6. Semantic Shifts Showing the Grammar and Structure of Bengali Comics and Japanese Manga



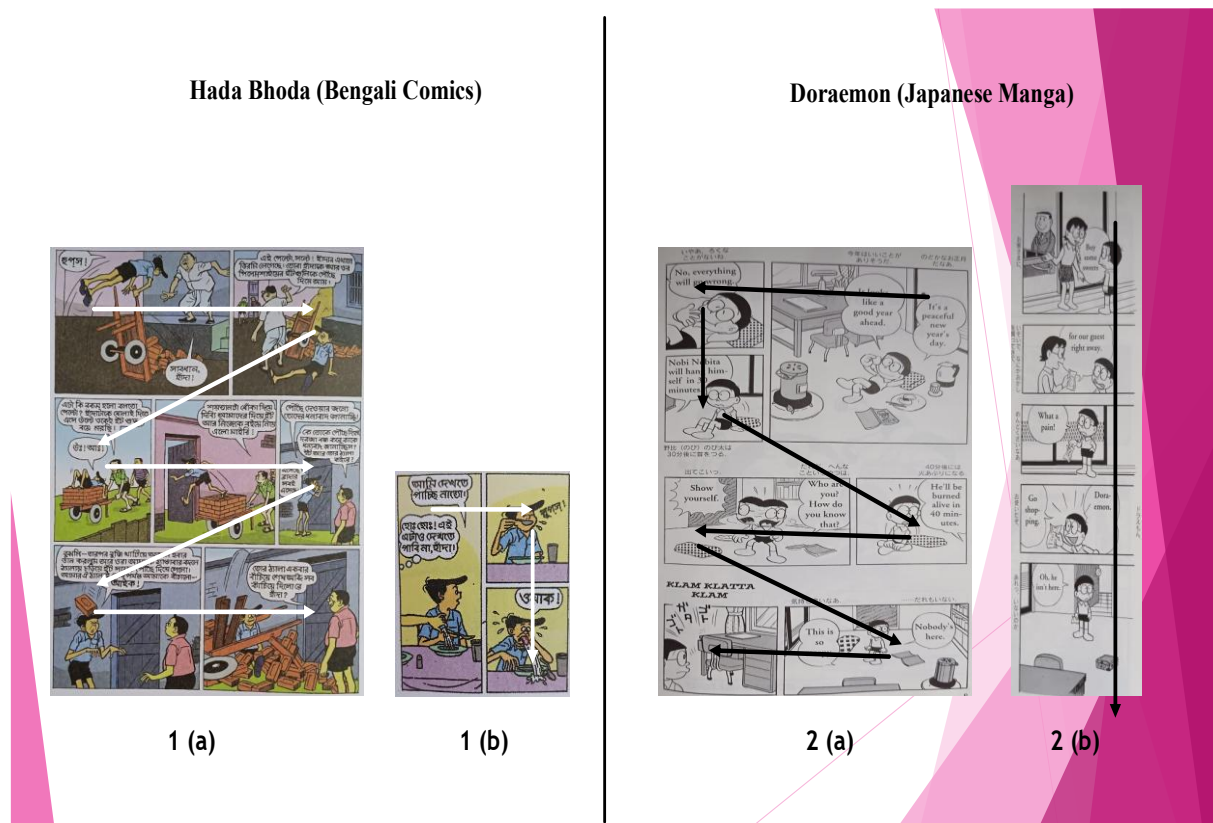
Source: Hada Bhoda and Doraemon series

The first part (left) of Figure 6. shows the two types of semantic shifts found in the Bengali comics, *Hada Bhoda*. 1 (a) shows a simpler way of the semantic shift from left to right and then again diagonally flowing from the right panel to the left panel of the second row; the same pattern is maintained throughout to form the whole narrative structure, following the common Western Z-path of reading comics. In addition, 1 (b) shows a different semantic

shift which goes from left to right and then goes vertically downwards. Similarly, on the other hand, the second part (right) of Figure 6., too, shows two different forms of semantic shifts in the Japanese manga, *Doraemon*. 2 (a) shows a rather complex semantic shift from one panel to another. It is read from right and then it takes a vertical flow and then diagonally flows to the right panel of the second row, where it is again read from right to left and continues with this pattern. Furthermore, 2 (b) shows a simpler form, which is read vertically, starting from the top panel to the bottom one. The empty spaces or the gutters between the panels found both in the Bengali comics and the Japanese manga play quite a significant role in shaping the flow of the narrative as the readers pursue reading; they also help with the transition from one panel to the other.

The semantic shifts showing the grammar and structure of *Hada Bhoda* and *Doraemon*, based on the skeletal structure shown in Figure 6., are as follows:

Figure 7. Semantic Shifts of *Hada Bhoda* and *Doraemon*



Source: *Hada Bhoda* and *Doraemon* series

As mentioned earlier, McCloud (1993) classifies the transitions between panels as the grammar of comics. These transitions depict how readers from different cultures interpret visual narratives as they shift their views from one image to the next. Each of McCloud's six transitions is examined in this study with the help of the two visual narratives, *Hada Bhoda* and *Doraemon*:

Figure 8. A Moment-to-Moment Transition



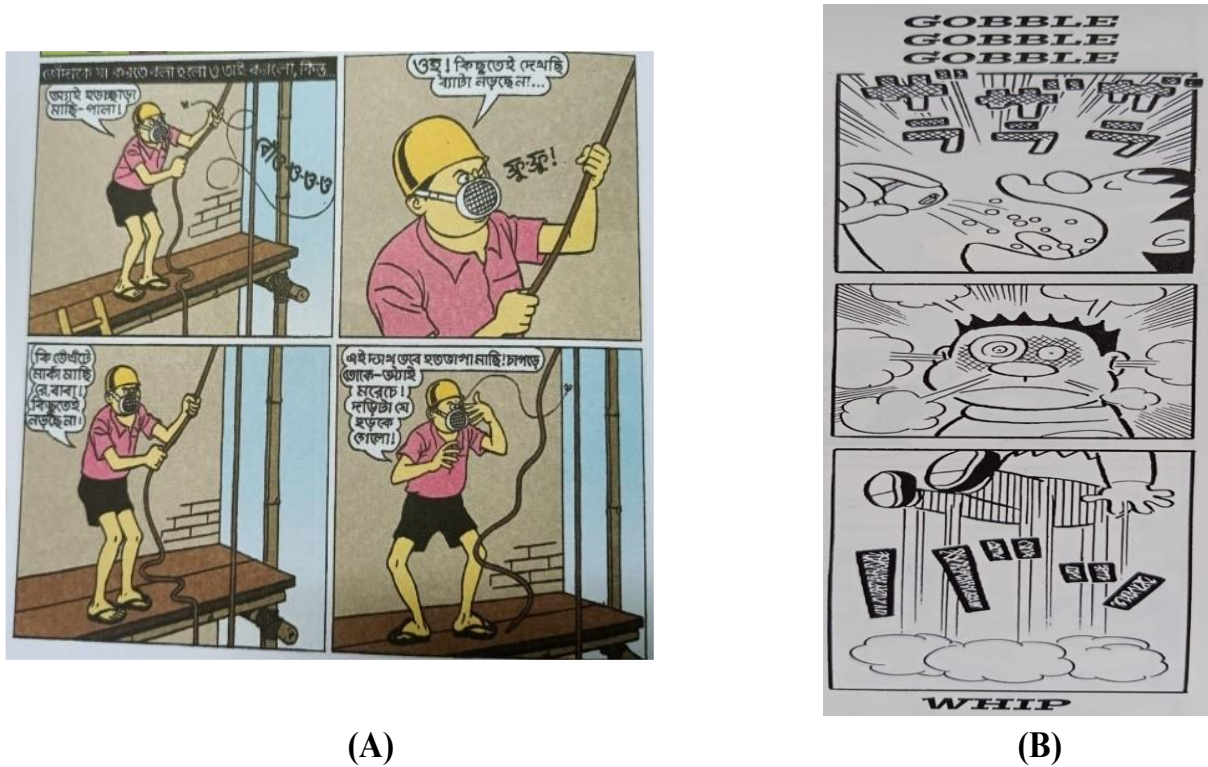
Source: Fujio's (1974) *Doraemon- Volume 1* (p. 52)-A; Debnath's (1962) *Hada Bhoda Samagra* (p. 9)-B

Moment-to-Moment: In this transition, very little variation exists between the two panels. It consists of a single action, which is portrayed in a series of moments—Figure 8.

The transition that occurs is the smallest here, with barely any time passing between panels, as if time were moving slowly. It is typically provided to employ in-depth attention to a definite narration or topic or to a specific area of that topic. A transition from one moment to the next could, for instance, concentrate on a face or a hand and then project a slight change in expression or gesture in the following panels, just like the above example from the *Doraemon* (A) series depicts. Similarly, the example from the *Hada Bhoda* (B) series depicts a scene that concentrates on the action where *Hada* tears off a necklace and the beads scatter like fallen stars throughout the three panels, and *Bhoda*, disgusted by this fact, wants to collect the beads. In both strips, (A) and (B), single actions are noticed throughout the

preceding panels, guiding the readers' cognitions as the sequences unfold. 'Moment-to-moment' refers to the continuous shift of communicative acts by representing certain gestures or building a conversation between speakers, respectively, present in the visual narratives.

Figure 9. An Action-to-Action Transition



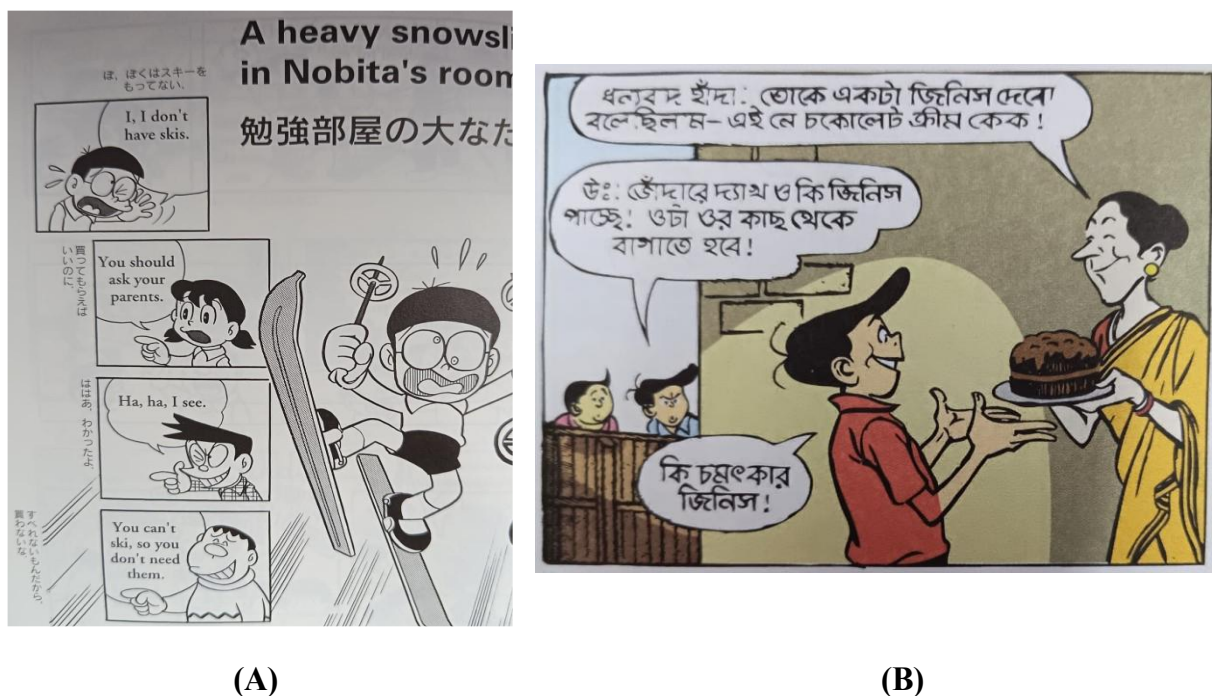
Source: Debnath's (1962) *Hada Bhoda Samagra* (p. 318)-A; Fujio's (1974) '*Doraemon*'- Volume 4 (p. 22)-B

Action-to-Action: In this transition, a single subject's action or a series of actions is displayed—Figure 9.

This is usually the most common panel-to-panel change that occurs in visual narratives. Even though there is just one subject in the centre of attention, time moves faster while the reader follows that subject as they go through a series of events. While the scene may shift a little, it will be one scene or location where the same subject performs many different acts or gestures. In Figure 9. (A), *Bhoda* has been given some work to do, but a fly keeps on disturbing him in different panels, depicting different scenes. Similarly, in (B), after taking too many of *Doraemon*'s magic pills, smoke billows from *Gian*'s nose and ears, and as a

result of it, he begins to soar too high, which is observed across the sequence of the vertically aligned panels. In both cases, a single subject's action is noticed throughout the panels. 'Action-to-action' is a sequential narrative approach in which the reader is allowed to follow a single subject through a sequence of consecutive acts or gestures. Even though the scene changes, the attention remains on the same subject who performs certain acts or gestures in the same scene, adding proper consistency and coherence to the narrative.

Figure 10. A Subject-to-Subject Transition



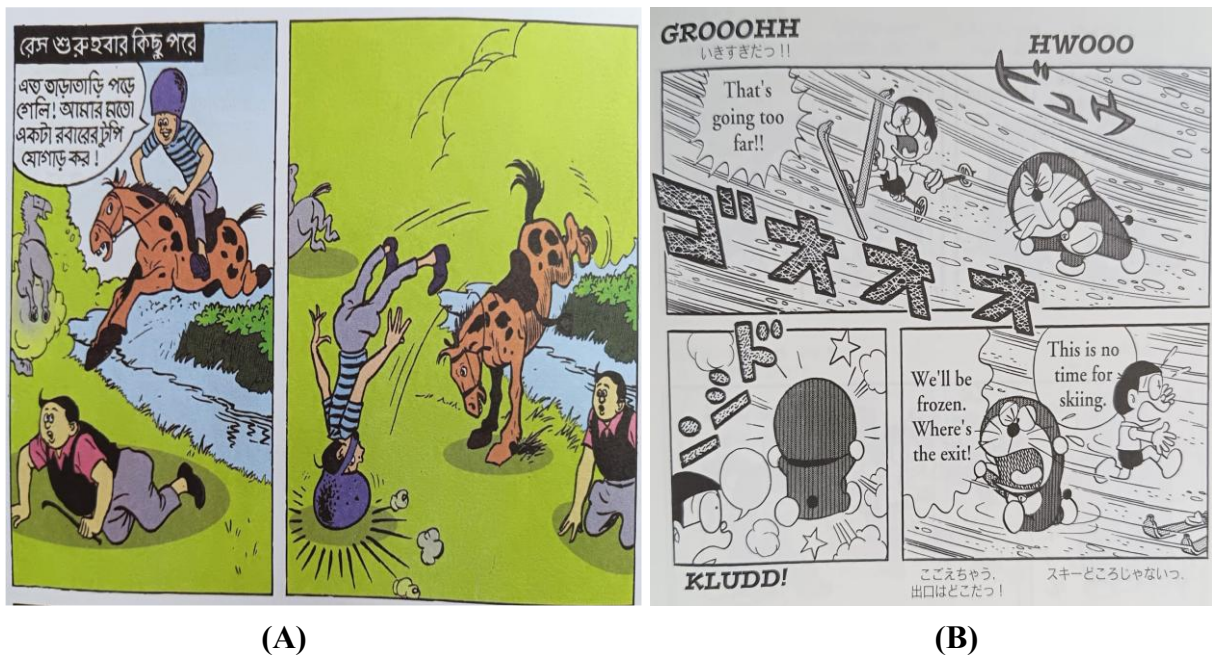
Source: Fujio's Doraemon- Volume 1 (1974, p. 83)-A; Debnath's Hada Bhoda Samagra (1962, 428)-B

Subject-to-Subject: It is a transition between various subjects in the same scene; it is a series of different changing subjects within a single scene—Figure 10.

The scene itself does not change while the reader experiences several subjects that make it up. It is a useful transition to show people in a conversation with their reactions to what is being said, or even to lay the ground for a definite setting. In Figure 10., such scenes are noticed from both series of *Doraemon* and *Hada Bhoda*. In (A), all the subjects are observed talking in the same scene about skiing while they are placed in vertically sequential panels.

Whereas, in (B), a single panel shows a lady presenting a beautifully decorated cake to *Hada's* doppelgänger, who is thrilled by the gesture, while *Hada* and *Bhoda* exchange sly glances as they think of snatching the cake away from him. ‘Subject-to-subject’ refers to the dynamic conversation of several speakers or subjects within a communicative context or scene. In scenes with multiple speakers, there are often complex exchanges.

Figure 11. A Scene-to-Scene Transition



Source: Debnath's (1962) *Hada Bhoda Samagra* (p. 24)-A; Fujio's (1974) *Doraemon- Volume 1* (p. 92)-B

Scene-to-Scene: It is the transition that takes us or transports us across significant distances of time and space—Figure 11.

This transition is not very common; it is useful in advancing the visual narrative without spending a lot of time on a specific narration or its details. The reader is allowed to assume that a certain event has occurred during the period that passes between a few panels. Nevertheless, there is a logical relationship established by the preceding panels. In (A), *Hada* mocks *Bhoda* for falling off his horse, boasting confidently about his rubber hat. In the very next panel, however, *Hada* is shown slipping off the horse himself. The event unfolds by

transporting the readers, relying on minimal text and the fluid movement of visuals across panels to guide readers through the sequence. In (B), it is observed that an avalanche is sweeping *Nobita* and *Doraemon* away, leaving them powerless as they're carried across the snow. The scene unfolds with minimal text, relying on explicit visuals to convey their struggle and lack of control to the readers. ‘Scene-to-scene’ refers to the transitions or changes between distinct communication circumstances or settings. It indicates that the readers can infer or presume that certain activities or exchanges occurred in specific scenes, even if they are not explicitly represented or written.

Figure 12. Aspect-to-Aspect Transition



Source: Fujio's (1974) *Doraemon- Volume 5* (p. 47)-A; Debnath's (1962) *Hada Bhoda Samagra* (p. 225)-B

Aspect-to-Aspect: It bypasses time by transitioning from one aspect of a place, idea or mood to another—Figure 12.

This is a non-temporal transition that can be employed in a dream sequence like in (A) or a telephonic conversation, like in (B) of Figure 12., respectively, as it depicts a single scene or subject in more than a single way. ‘Aspect-to-aspect’ refers to the analysis or portrayal of a particular situation from multiple perspectives or dimensions during a discourse. This is a common notion, used to strengthen the reader's understanding of a specific scenario by

investigating several features. In the context of a dream (sometimes also fantasy) sequence or a telephonic conversation, it entails evaluating the discourse from multiple angles, such as perspectives, tones, emotions, non-verbal cues, etc.

Figure 13. A Non-Sequitur Transition



Source: Debnath' (1962) *Hada Bhoda Samagra* (p. 488)-A; Fujio's (1974) *Doraemon- Volume 5* (p. 14)-B

Non-sequitur: This transition provides no logical relationship between panels—Figure 13.

In both (A) and (B) of Figure 13., no logical relationship is found between the panels, whether read left to right for *Hada Bhoda* or right to left for *Doraemon*. Nevertheless, it must be kept in mind that although there do not seem to be any connections between the panels right at the moment, they might develop as the storylines or narrations proceed further. 'Non-sequitur' refers to a statement or utterance that does not follow any logic or a coherent progression from the preceding context or conversation.

Therefore, these transitions shape the rhythms, tones, and reader interactions, allowing visual narratives to convey complex storytelling techniques that smoothly blend visual and textual elements. McCloud's (1993) framework also emphasises how comics use a unique 'visual grammar' to move readers through the stories without relying solely on traditional narrative structures.

4.3 Functional Aspects of Visual Narratives—Society in Miniature

Moving beyond the structural and grammatical aspects of visual narratives, this study found that a major part of these visual narratives depends on their functional aspects. The linguistic standards and expectations of a society are shaped by the functional aspects of the visual narratives. This research shows the investigation of how bimodal communication in visual narratives conveys the functions present in them through family dynamics, friendly exchanges, school conversations, sportsmanship and culinary culture:

4.3.1 Family Dynamics

In Figure 14., the vocabulary used in Bengali reflects the colloquial Bengali speech spoken in everyday life. Here, *Hada* is asking his fraternal uncle, ‘How do you like the wheelbarrow, uncle? I made it myself.’ This is a declarative utterance that is providing information or expressing a proposition. In this scenario, the speaker, *Hada*, is informing his uncle about the wheelbarrow and its provenance (which he created himself). While the utterance also includes a

Figure 14. Debnath's (1962) *Hada Bhoda Samagra* (p. 61)



question (How do you like...?), the sentence's general structure and intent are more focused on presenting information than eliciting information from the listener. In response, his uncle is saying, ‘It is wonderful!’ This shows an evaluative utterance that expresses an assessment or judgment about the wheelbarrow; the uncle is expressing his view on the speaker’s wheelbarrow. By saying ‘it is wonderful,’ the uncle expresses his appreciation for the wheelbarrow. The uncle's pleasant and positive answer displays familial support and encouragement. In many cultures similar to this one, family members often promote and celebrate each other's accomplishments, creating a supportive environment in which people feel valued and driven to follow their passions and abilities. The uncle expresses respect for

Hada's wheelbarrow, recognising the effort and skill spent in its production. The uncle's gratitude for the wheelbarrow deepens the bondage between the two family members; the happy and proud facial expressions add to it. Appreciation is a critical component of emotional expression in families. The uncle's affirmative remark not only communicates his opinion about the wheelbarrow but also expresses affection and warmth for *Hada*, establishing emotional links and familial ties. The visual narrative shows that in Bengali culture, respect for handmade objects or craftsmanship is firmly ingrained. By expressing respect for the wheelbarrow, the uncle may be passing down cultural values such as craftsmanship, ingenuity, and self-reliance to future generations, thereby sustaining family traditions.

Figure 15. Debnath's (1962) *Hada Bhoda Samagra* (p. 490)



In Figure 15., *Hada* and *Bhoda's* aunt brings food for *Bhoda* and says, '*Bhoda*, eat! I have made a special meal for you as per your desire.' It is an imperative utterance that directs or encourages *Bhoda* to take a specific action. In this scenario, the speaker, the aunt, instructs the listener, *Bhoda*, to eat while also notifying him that she has prepared a special meal based on *Bhoda's* preferences. The imperative 'eat' denotes a directive or command, whereas the

statement about the special meal provides extra information or motivation for *Bhoda* to follow the demand. Even though, by the facial expression of the brothers, it seems that *Hada* has planned some mischief against *Bhoda*, the love that the aunt has for *Bhoda* is unconditional. The amount of food that the aunt is serving to her nephew also implies that in Bengali culture, serving food to family members is very important, both culturally and emotionally. It can be called a type of 'love language' because it emphasises caring,

affection, and familial relationships. Serving special Bengali dishes to family members not only honours culinary heritage, but also fosters cultural identity and pride. Also, it is observed here that the aunt used the English word ‘special’ in between her Bengali dialogue. This can be noted as an instance of a Bengali family in the post-colonial era. *Hada Bhoda* was created after the independence of India; a culture can be an imagined community, but it reflects multi-cultural influences. Bengali cultural identity contains two distinct elements: precolonial Bengali presence and British European presence (Zahin & Nion, 2023).

In Figure 16., *Bhoda* is found to be scolded by his father as he says, ‘Rascal! You’re cleaning the drain with a vacuum cleaner! Clean it soon!’ This is an example of a declarative phrase followed by an imperative statement. The first section of the utterance is a remark that provides information on *Bhoda*’s action or behaviour. Declarative utterances merely state or affirm something without requiring

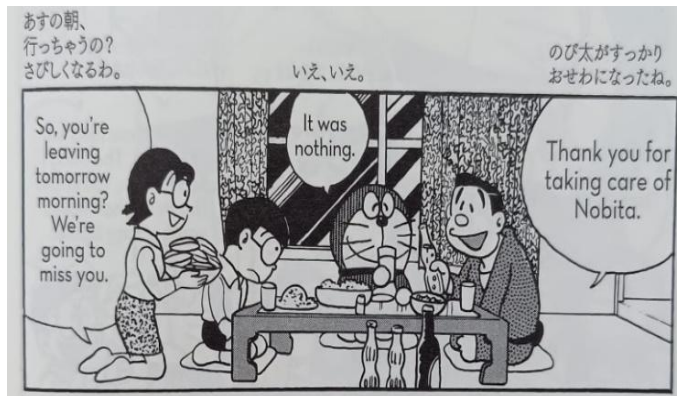
Figure 16. Debnath’s (1962) *Hada Bhoda Samagra* (p. 392)



the listener to perform any specified action. While ‘Clean it soon!’ is an order or directive, instructing *Bhoda* (referred to as ‘rascal’) to clear the drain right away. Imperatives are distinguished by the function of issuing commands or making requests. *Bhoda*, on the other hand, is found to be screaming, ‘Dad!’ as he is getting pulled by the ears. Whereas, *Hada*, his mischievous brother, is noticed hiding and laughing at *Bhoda* from a distance. The father’s utterance is used together to communicate unhappiness or disapproval of *Bhoda*’s behaviour (cleaning the drain with a vacuum cleaner) while also providing direction or instruction on what needs to be done. It implies that in Bengali culture, parental scolding can be severe, but it is often motivated by a desire to inculcate discipline, direction, and values in children. The angry facial expression of the father implies dissatisfaction, while *Bhoda*’s

expression reveals pain and agony. On the other hand, *Hada's* laughter implies sibling-bullying in a playful, teasing manner.

Figure 17. Fujio's Doraemon- A Selection of Touching Stories (p. 042)



In Figure 17., it is found that the Japanese family is dining together while having a heartfelt conversation. It indicates that *Doraemon* is leaving the family soon, and they are having their last supper together. The mother saying, 'So you're leaving tomorrow

morning? We're going to miss you.', starts with an interrogation followed by a declarative sentence that expresses a declaration or fact. Specifically, it conveys a farewell sentiment and expresses the mother's sadness or regret about *Doraemon's* upcoming departure. But she bids farewell with a smile on her face anyway. Like the Bengali culture that is observed in Figure 14., serving food to a family member (even though it is a robot here) is a love language as it emphasises caring, affection and familial relationships. Here, in this figure, she is observed to be serving *Doraemon's* favourite food, *dorayaki*, while the father thanks *Doraemon* for taking care of his son, *Nobita*. Thanking *Doraemon* before leaving is an expression of gratitude and gratefulness. It acknowledges his generosity, hospitality, or assistance, and it is a method to express gratitude, which is also clearly reflected in the facial expression of the father. This gesture is generally interpreted as polite and respectful, particularly in social or formal settings. It also displays affection and acceptance for a robot, meaning that despite its non-human nature, *Doraemon* has been treated with care and tenderness comparable to that of a family member. Acceptance of outsiders is often regarded as an expression of love in Japanese society, particularly in familial and community settings. This notion is based on the Japanese virtue of *omotenashi*, which entails hospitality, kindness, and care for others, particularly those who may be considered strangers; it is the Japanese service quality, culture,

and background that are mostly derived from the heart (Belal et al., 2012). *Doraemon's* response with 'it's nothing,' which is a prevalent cultural response in many societies, demonstrates humility and modesty on the part of the mother and the father, who offered food and felt gratitude, respectively. On the other hand, *Nobita's* facial expression implies sadness because his brother-like-best-friend, *Doraemon*, is leaving. The melancholy that comes with someone's leaving is a normal emotional response to the impending separation of a family member.

Figure 18., depicts a scene that is deeply emotional and passionate, eliciting sentiments of relief, joy, and overpowering love. The mother is found crying and hugging *Nobita* tightly while she says, 'I thought you'd died. You were just lying there without breathing.' ... 'What a relief!' The utterance is a compound one as it consists of two independent clauses connected by a hidden coordinating

Figure 18. Fujio's *Doraemon- Volume 4* (p. 140)



conjunction 'but' in place of the full stop. It is a declarative sentence followed by an exclamatory utterance displaying intense emotions or feelings, such as relief and surprise. Every heartbeat of the mother seems like a symphony of delight, and every breath is a prayer of gratitude from the visual interpretation. On the other hand, *Nobita* looks a little confused and clueless, yet he seems equally moved too. It depicts the unbreakable relationship between mother and child while reminding us of the value of life.

Figure 19. Fujio's *Doraemon- Volume 2* (p. 28) (Read: Right to Left)



In Figure 19., *Nobita* is found to be scolded by his mother, similar to what we found in Figure 16. from the *Hada Bhoda* series. Here, ‘Since this morning you’ve been saying ‘later, later’, but you haven’t studied at all.’, is a compound sentence; it consists of two independent sentences (‘Since this morning you’ve been saying “later, later”’ and ‘you haven’t studied at all’) connected by the coordinating conjunction ‘but,’. It is also declarative as it presents a statement of fact. The English language typically avoids word reduplication, but since this phrase is translated from Japanese, ‘later, later’ appears in the English version too; that is why it is kept in quotes as well. The statement ‘Since when do you hold a pencil with your left hand? Do it right.’ follows the former sentence, which starts with an interrogative sentence that sounds more like a self-assurance question, followed by an imperative one as it directs *Nobita* to perform a specific action. This figure is a complete contrast to that of the previous one (Figure 18.) as the mother’s facial expression is full of rage and she is pulling *Nobita*’s shirt to make him study. Getting scolded by parents seems to be a typical approach in both the Bengali and Japanese cultures. They offer advice on how to improve the issue and encourage their children to consider their actions, which is observed to be slightly harsh in both cultures, as noticed through the visual narratives. Here in this figure, *Nobita*’s facial expressions first depict disinterest and later tearfulness as he is forced to study.

4.3.2 Friendly Exchanges

In Figure 20., it is observed that *Hada* is feasting with his friends. The boy in the orange shirt is saying, ‘Keep eating happily, friends! My *Nyapa*-uncle is the chief waiter here.’ which starts with an imperative sentence as it gives a command or instructs the friends, encouraging them to continue eating happily. The latter part of the utterance is a declarative sentence. It makes a

Figure 20. Debnath's (1962) *Hada Bhoda Samagra* (p. 287)



comment or provides information on the speaker's *Nyapa*-uncle, who is the chief waiter at the restaurant. The utterance refers to a communal or social dining experience in which people are urged to enjoy themselves in the company of others. This reflects cultural ideals such as camaraderie, hospitality, and conviviality, in which shared meals facilitate bonding and connection among a community or social group. The phrase ‘My *Nyapa*-uncle is the chief waiter here’ appears to indicate a familial or personal relationship with someone in a position of authority or service in the context (even though it is just the chief waiter here). This implies a cultural norm or practice of nepotism or familial influence in particular situations when family ties or relationships may have an impact on one's function or standing within a profession or institution. The facial expressions of the friends add to it—they eat the food without hesitation as if they are family. It is already discussed in Figure 15., how food can act as a ‘love language’ in Bengali culture. Similarly, here, food is strongly connected to cultural customs and symbols. Sharing food with friends is often connected with hospitality, warmth, and acceptance, all of which are signs of love and inclusion.

Figure 21. Debnath's (1962) *Hada Bhoda Samagra* (p. 462)



In Figure 21., it is observed that *Bhoda* is getting bullied by *Hada* and other friends. *Hada* is saying, 'You can enjoy the treat from outside, *Bhoda*!' which is an imperative statement giving a command, or rather instructing *Bhoda* to enjoy the cake from outside. It is a birthday party where *Bhoda* is tricked and bullied by his friends and excluded from the celebration. Humour is used here to diffuse the situation while also communicating the message that exclusion is unacceptable. *Bhoda* has an angry facial expression, while his friends' faces look quite happy while bullying him.

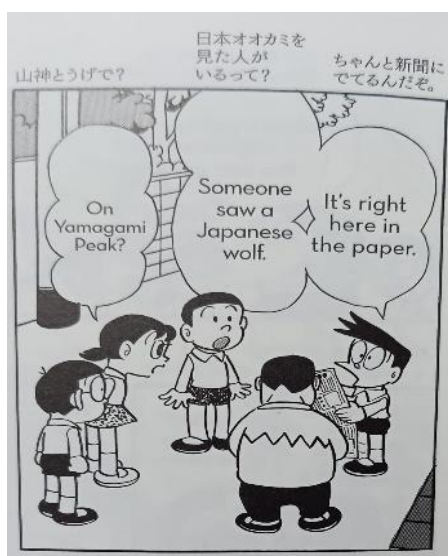
Figure 22. is a continuation of the previous panel, Figure 21. In Figure 21., it is already observed that *Hada* tried to be sarcastic by saying that *Bhoda* could enjoy the cake from outside. *Bhoda* here, acts smart and counterattacks *Hada* by saying, 'So you guys thought this!' while he takes away the cake with a long stick. *Hada*, along with other friends, on the other hand, is confused due to this act, and one of the friends

Figure 22. Debnath's (1962) *Hada Bhoda Samagra* (p. 462)



says, 'What's this! Our cake is going away!' *Bhoda*, being bullied by his friends, decides to play a harmless trick on their pals in retaliation. Due to this, the situation became more humorous and lighter, leaving the bullies confused and clueless. From this, it can be inferred that pranking among friends is a fun way to socialise and share laughter in the Bengali culture.

Figure 23. Fujio's (1974) *Doraemon- A Selection of Touching Stories* (p. 005)



In Figure 23., *Nobita* and his friends gather in a place to discuss a quite serious issue. *Suneo* has a newspaper in his hand, and he talks about something that he read in it: ‘Someone saw a Japanese wolf.’ It is a declarative sentence as it presents a fact. These friends are passionately discussing Japanese wolves, lending an air of excitement and amazement to their conversation, delving into the fascinating realm of this extinct animal and picturing its presence in Japan's wilds. Japanese wolves, also known as *Honshu* wolves, were a subspecies of grey wolf indigenous to Japan. They were declared extinct in the wild by the early twentieth century. Yet, many of the reported sightings of the Japanese wolf have happened in the *Kii* Peninsula, in the south-central region of the main island of *Honshu*; they are also called *yamainu*, meaning *mountain dog* (Knight, 1997). *Shizuka*’s facial expression, along with others, seems to be in awe as she talks about *Yamagami* peak, which is probably a fictitious mountain that means *mountain god* (Shiko, 1936); it can also probably be Mount *Kami*. It seems that in Japanese culture, kids like to gather around to seriously discuss mythical and fictitious figures, which can be an entertaining and educational activity for children.

In Figure 24., *Nobita* is getting bullied by *Gian* and *Suneo* just like *Bhoda* was getting bullied in Figure 21. His yo-yo is taken away by *Gian*. *Suneo* asks sarcastically, ‘Wouldn’t the yo-yo be happier

Figure 24. Fujio's (1974) *Doraemon- Volume 4* (p. 110) (Read: Right to Left)



being played with by *Gian*?' which is an interrogative sentence where the 'yo-yo' is personified. The utterance is sarcastic since it implies that the yo-yo would be happy in *Gian*'s ownership, even though it originally belonged to *Nobita* and was taken away from him. This sentence can be categorised as an ironic or sarcastic interrogative sentence. It implies a rhetorical query, conveying *Suneo*'s disapproval of the circumstance while emphasising the absurdity or injustice of the yo-yo being removed while *Gian* enjoys and speaks out an exclamatory utterance in support: 'Good idea!' He further says, 'If you get better at the yo-yo, I'll give it back to you.' This utterance is both conditional and sarcastic. It uses the conditional clause 'If you get better at the yo-yo', to give the impression that *Nobita* can retake ownership of the yo-yo. However, *Gian* has no intention of returning the yo-yo, and he uses sarcasm to criticise *Nobita*'s lack of ability with it. On the other hand, *Nobita* is found to be crying heavily at this whole situation. Therefore, it can be said that bullying exists among friends in the Japanese culture as it does in the Bengali culture.

4.3.3 School Conversations

Figure 25. Debnath's (1962) *Hada Bhoda Samagra* (p. 123)



Conversations in school include performative acts

where individuals utilise language to accomplish tasks. Here in Figure 25., the teacher is found to be teaching about 'Androcles and the lion'. He says, 'According to an old Greek story, when Androcles saw a thorn in the lion's foot, he plucked it out with great compassion.' It is a declarative sentence since it states a fact about the story and explains an action

made by Androcles when he noticed a thorn in the lion's foot. The introduction of Greek mythology in Bengali classrooms can be interpreted as a type of cultural interchange. Greek mythology is culturally and historically vital in Western civilisation, and its incorporation within Bengali education exposes students to other cultural perspectives and narratives from

throughout the world. The formal study of literature as a body of knowledge began in India with the establishment of modern colleges and universities in the early and mid-nineteenth centuries. Following William Bentinck's English Education Act of 1835, English literature became a part of liberal education in India (Alexander, 2023). It is also inferred that the fable of Androcles and the Lion is often presented in Bengali classrooms to teach children about compassion. Androcles' interests were not served by such a good act in the woods, where no one could see it, reward it, or prevent the lion's predatory or callous response. However, Androcles' sympathetic gesture of serving the lion's interests earned the lion's trust (Bell, 1998).

In Figure 26., the teacher seems quite angry with *Bhoda*. He says, '*Bhoda* is late today as well! He arrived late to school seven days in a row. If you are late again, you will not be permitted to enter the school.' The first two sentences of the utterance are declarative, but the last part makes the utterance an imperative one as it is a command or instruction directed towards *Bhoda*. The cultural practice depicted in this scenario reflects classroom norms and

Figure 26. Debnath's (1962) *Hada Bhoda Samagra* (p. 125)



expectations, particularly regarding punctuality and discipline. The teacher has authority and is in charge of keeping discipline in the classroom. The teacher sets behavioural limits and expectations by establishing clear rules and consequences for tardiness. Also, by looking at *Bhoda's* expression, it can be well understood that he is guilty of not being on time for class.

Figure 28. Debnath's (1962) *Hada Bhoda Samagra* (p. 50)



Figure 27., depicts that music is also included in school curricula alongside traditional academic topics. Here, *Hada* is complaining about the instrument that he has been given to play. He is saying, ‘Oooooooo! My talent is being lost with this instrument that sounds like “ting ting”.’ which is a declarative sentence. Here, some Bengali onomatopoeic words are found to be uttered by him, which mimic natural sounds that differ significantly from other languages: (a) *Oooooooo*—A complaining expression. (b) *Ting ting*—the sound of the instrument that *Hada* is playing. This participation of *Hada*, who is unable to play any instrument, also implies that the activity or event was not exclusively focused on musical skill or proficiency, but also on fostering a sense of belonging or fellowship among group members. But the writer still preferred him to continue to criticise out of pride, as well as to keep the comics interesting and entertaining.

Here in Figure 28., the teacher is found to be motivating *Nobita*. He is asking him not to dwell on the past and advising him to stay focused on the future, which are both imperative utterances. On the other hand, *Nobita*'s utterance is an affirmative response indicating agreement. His facial expression

looks quite emotional as he gets motivated. The teacher's function as a motivator reflects a cultural respect for education and a belief in its transforming ability for individuals and society. This suggests that the Japanese society promotes humility and continuous self-improvement.

Figure 27., depicts that music is also included in school curricula alongside traditional academic topics. Here, *Hada* is complaining about the instrument that he has been given to play. He is saying, ‘Oooooooo! My talent is being lost with this instrument that sounds like “ting ting”.’ which is a declarative sentence.

Here, some Bengali onomatopoeic words are found to be uttered by him, which mimic natural sounds that differ significantly from other languages: (a) *Oooooooo*—A complaining expression. (b) *Ting ting*—the sound of the instrument that *Hada* is playing. This participation of *Hada*, who is unable to play any instrument, also implies that the activity or event was not exclusively focused on musical skill or proficiency, but also on fostering a sense of belonging or fellowship among group members. But the writer still preferred him to continue to criticise out of pride, as well as to keep the comics interesting and entertaining.

Figure 27. Fujio's (1974) *Doraemon- Volume 3* (p. 3)

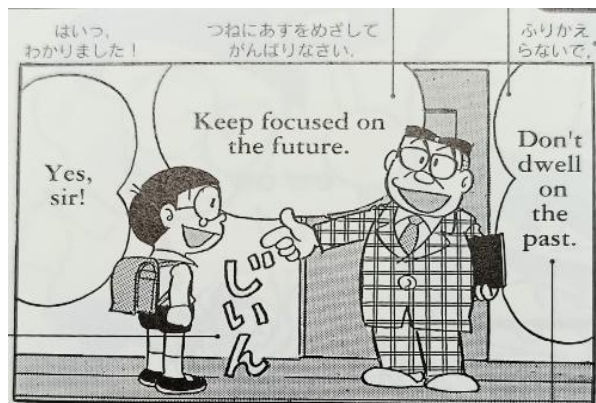


Figure 30. Fujio's (1974) *Doraemon- Volume 4* (p. 120)



Figure 29. is somewhat similar to the Bengali classroom culture present in Figure 26. The teacher is equally upset (like Figure 26.) since *Gian* and *Suneo* were late for an additional time in a row. *Gian* and *Suneo*'s words imply a cultural environment in which punctuality is valued, and there may be a sense of rivalry or comparison among individuals in terms of adherence to timetables or standards. *Nobita*, on the other hand, appears content with a happy facial expression as he maintains his punctuality.

It can be inferred from Figure 30. that the Japanese culture highly values academic achievement and excellence. The teacher's expression of disappointment and his utterance that the test was difficult and nobody scored well may stem from cultural

Figure 29. Fujio's (1974) *Doraemon- Volume 1* (p. 53)



expectations of students' academic performances. On the other hand, *Gian* speaks out of pride as he has scored full marks, while others look confused and are bewildered about how he obtained the marks. This also fosters a sense of competitiveness among students, as discussed in Figure 29.

4.3.4 Sportsmanship

Figure 31. Debnath's (1962) *Hada Bhoda Samagra* (p. 16)



Figure 31. shows *Hada* kicking a ball with great force, causing it to hit and destroy a neighbouring window or glass panel. He is saying, 'Have you felt the power of my shot? I cracked the glass even with this ball.' which starts with an interrogation and is later followed by a declarative statement as it talks about an action that has already occurred. He emphasised by saying, 'even with this ball' as it is not quite a strong one, that it can break the

glass. It seems his pride in this moment extends beyond mere satisfaction to a deep sense of accomplishment that resonates within him, leaving the other boy surprised. This depicts that football is a game that is intertwined with Bengali culture: Bengalis are well-known for their love of football. It maintains a particular place in the hearts of Bengalis, having evolved from a basic sport into a cultural phenomenon. They incorporated football into their traditional activities, and it became the 'pathological game' of Bengalis (Nasim, 2019). Also, the visual narrative comes with an onomatopoeia that creates a sound like *jhonnn*, which is a culture-specific sound of breaking glasses typical to the Bengali language.

The visual narrative in Figure 32. captures the thrill and activity of a cricket match by spotlighting crucial events on the pitch and demonstrating the emotions of the players. The visual cues of the spinning of the bat, the moving direction (motion lines) of the ball, and the bodily gestures add to it with multisensory effects. *Bhoda* appears to be

Figure 32. Debnath's (1962) *Hada Bhoda Samagra* (p. 18)



bowling, and the boy next to him exclaims, ‘See how well *Bhoda* has bowled.’ which is a declarative sentence. The wicketkeeper is shouting, ‘Out!’ in mounting excitement, making it an exclamatory utterance, and the batsman appears dejected as he stammers and responds, ‘How did this happen?’ which is again an interrogation. Cricket in India is more than simply a game; it is a national passion that has been ingrained in Bengali culture. It was introduced to India by the British during the colonial period. Most Indians have always considered cricket to be more than just a game. The sport infuses India's fundamental fabric with social subtexts that extend beyond the cultural, political, and religious (Nair, 2011).

Figure 33. Fujio's (1974) *Doraemon- Volume 5* (p. 52)



Figure 33. depicts *Nobita* and *Doraemon* using *Doraemon's* gadget to travel back in time. They have met *Nobita's* father and grandfather in the past. The visual narrative shows a traditional Japanese martial art, *kendo*, which

is focused on the use of the Japanese sword. *Kendo* literally means ‘the way of the sword’, but Japanese swordsmanship was previously known as *kenpo* and *kenjutsu*. *Kendo* was a discipline that developed a *samurai* personality during the *Tokugawa* period and played an essential part in national education (Sakaue, 2018). Even though *Nobita's* father is having a hard time keeping up with the sport, *Nobita's* grandfather looks quite strict about it and says not to give up with an exclamatory tone. Thus, it is clear from the narrative that *kenjutsu* training emphasises discipline, focus, and control over both the body and the mind. According to Sakaue, Japanese martial arts are the most important as a means of comprehending the pure essence of Japan, and are ideal for creating the Japanese people's awareness.

Figure 34. shows a baseball match scenario where *Nobita* and his friends are found to be playing. It illustrates how *Nobita* spins around and hits the ball, leaving everyone in awe and delight. Even *Nobita* shouts out with excitement that it is unbelievable to him, followed by a declarative utterance that he has hit the ball. Baseball in Japan is quite a popular sport, even though it is not native. Baseball's tremendous appeal in Japan is due in part to the fact that the sport originated with high school and college teams rather than a professional club system (Ikei, 2000).

Figure 34. Fujio's (1974) *Doraemon- Volume 1* (p. 133) (Read: Right to Left)



4.4.5 Culinary Culture

Figure 35. Debnath's (1962) *Hada Bhoda Samagra* (p. 71)



Figure 35. depicts a scenario where *Hada* and *Bhoda's* aunt is frying fish while a ghost is seen stealing them from behind. The aunt, unaware of the ghost, says to herself, 'Let me fry them, we'll dine together later.' which is a declarative statement of her present action and the action that she has predicted for the future. On the other hand, the ghost has quite a long hand to get hold of the fried fish: Bengali literature and culture are heavily

reliant on ghosts. There are quite a few kinds of ghosts found in Bengali literature—among them, *mechho bhoot* are the ghosts that are fascinated with fish and can 'beg, borrow, steal or kill' for a fresh slice of fish, especially if it is hilsa (Maiti, 2015). Moreover, Bengali people are very fond of fish, and fish is not only the most commonly consumed protein but is also the preferred one (Akter, et al, 2019). Therefore, it is also found in the Bengali comics, *Hada*

Bhoda, that fish plays a very significant role in Bengali cuisine. However, the geographical location of West Bengal, bordered by rivers and the Bay of Bengal, provides numerous options for fishing (Dan, 1985). This figure also illustrates its multisensory effect of how the fishy smell is rattling away with the wind, and as a result, it makes the ghost unable to resist himself. Also, it is prominent from the visual cues that eating with hands is also embedded in the Bengali culture.

Figure 36. shows a scene where a man is getting *biryani* for *Hada* and *Bhoda*. The man says, ‘Here is your *biryani*; eat comfortably. If you want more, call me, I am inside.’, which is an imperative utterance as it is an instruction. *Hada*, on the other hand, is delighted to see the *biryani* and tells *Bhoda*, who is skeptical about the man, ‘See, *biryani* has come and still you’re wondering about

Figure 36. Debnath’s (1962) *Hada Bhoda Samagra* (p. 71)



the way he stares.’, which is a declarative sentence. Even though they were suspicious about the man, seeing ‘*biryani*’, they could not resist their hunger. *Mughlai* food was introduced in Bengal under Mughal emperor Shah Jehan in the 17th century, when Burdwan became a revenue collection centre for Bengal (Sunder, 2020). *Biryani* and other *Mughlai* cuisine, with its rich flavours and aromatic spices, has left an imprint on the gastronomic scene of West Bengal, particularly in Kolkata and surrounding districts.

Figure 38. Debnath's (1962) *Hada Bhoda Samagra* (p. 71)



In Figure 37., *Bhoda* is too happy with his sweet and savory dishes, and he exclaims with joy, ‘Absolutely delicious! I like chop-cutlets as much as I like apples made of *kheer*.’, which starts with an exclamatory followed by a declarative utterance. The man in front is also happy to serve as he says, ‘Eat, eat *Bhoda*! Eat as you wish!’ which is an imperative utterance.

On the other hand, *Hada* seems disappointed as he is missing out on the delicious delicacies. Bengal is well-known for its extensive and diverse selection of sweets, known as *mishti*. Here in this narrative, *Bhoda* is having apples of *kheer*, which is a sweet; these sweets are a significant part of Bengali culture and are consumed throughout a variety of festivals and ordinary events. *Mishti*, regardless of caste, creed, or culture, plays a vital role in every Bengali household (Sarkar, et al, 2023). In addition, *chop-cutlets* are famous street snacks in Bengal, which are sometimes sold alongside *mishti* in sweetshops. These are generally spiced, round or oval potato cakes filled with ground fish, meat or vegetables, dipped in eggs and breadcrumbs (Sen, 1996). Also, *Bhoda* is found to eat with a fork and a knife here, which is again an influence of post-colonial India (West Bengal).

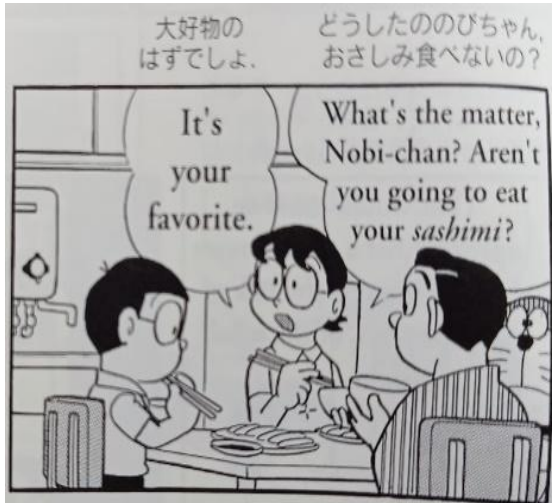
Figure 38. depicts that fish is an important aspect of Japanese cuisine, just like Bengali cuisine, as observed in Figure 35., and Japan is well-known for its seafood-centric meals. By looking at the face of the man when he is interrogated by *Nobita* about the fish he is holding, it can be inferred

Figure 37. Fujio's (1974) *Doraemon- Volume 5* (p. 22)



how much fish is liked by the Japanese people. The man declares that he will either grill it or fry it: a dish of charcoal-grilled fish seasoned with salt frequently features as the centrepiece in Japanese meals (Gurunavi, 2017).

Figure 39. Fujio's (1974) *Doraemon- Volume 5* (p. 22)



In Figure 39. *Nobita's* mother is found to interrogate him regarding his ignorance of his favourite food, *sashimi*, and others look shocked too. *Nobita*, on the other hand, looks like he is thinking about something else. The Japanese diet traditionally relied mainly on rice for calories and on seafood for animal protein (Swartz, 2004); it is evident from the visual cues, too. *Nobita*, too, is found to be fond of *sashimi* or

raw fish, which has been eaten in Japan since 500 BCE and is one of the most popular ways to enjoy fresh fish in Japan. This fish is cut neatly and presented on a tray (as depicted in the visual narrative) and served with soy sauce and condiments such as fresh *wasabi*, grated ginger, and shredded *daikon* radish (Gurunavi, 2017). Also, *Nobita* and his family are found to be using chopsticks from the visual cues, which represent the traditional way of eating Japanese food.

Figure 40., shows happy faces as the lady declares that she has made some *ohagi* for their guest, *Nobita's* father; even though making *ohagi* costs time and work, *Nobita's* father acknowledges it and indicates that it was not necessary; a subtle humbleness is conveyed through this gesture. The characters are also found to be seated in the

Figure 40. Fujio's (1974) *Doraemon- Volume 4* (p. 57)



traditional Japanese manner. The traditional *wagashi* (confectionery) from Japan combines local ingredients as well as local cultural elements, thus localised *wagashi* (He, 2018). These Japanese cakes and sweets provide a fascinating variety of flavours, textures, and creative displays; *ohagi* is a *wagashi* made out of rice flour, bean paste and sugar.

From the above examples, it can be inferred that grammar, structures, and functions go hand in hand even in visual narratives, just like spoken and written languages. While grammar and structures provide the frameworks for the visual language, their functions define the purposes. The arrangement of images and texts together allows the readers to convey specific meanings, intentions and relationships (function).

4.4 Discussion

Notably, both Chomsky's (1957) TGL and Halliday's (1961) SFL are present in visual narratives. By analysing the two visual narratives, *Hada Bhoda* and *Doraemon* from two different cultures, Bengal and Japan, respectively, it is found that structures and grammar serve as the foundation for storytelling, while the functions of these aspects bring it to life. The structure of a visual narrative refers to the physical layout, which includes panels, gutters, and overall page composition. The grammar includes the visual language and the specific rules that help in arranging the panel sequences. The functions, on the other hand, are concerned with how these structural and grammatical aspects work together to convey meanings, evoke emotions, and communicate culture-specific values. Furthermore, the notion that visual languages, too, have structures involves readers learning and decoding visual narratives through their specific structures and grammar. Just as children acquire language while they observe their surroundings and imitate the use of words in context, readers of visual narratives, too, develop interpretative skills by engaging with the images and texts present in them. They learn through visual cues with specific structures and grammar that eventually convey functions, associating them with social norms and values, thereby allowing for deeper levels of understanding of the picture stories. The social aspects surrounding these

visual narratives significantly reflect the cultural, historical, and social contexts in which they are created, serving as mirrors of the society and capturing the values, norms, and issues relevant to culture-specific communities. For instance, the portrayal of characters, settings, and culture-specific topics in *Hada Bhoda* and *Doraemon* reveals significant insights into social and cultural identities. Therefore, the readers explore complex themes through the communal engagement that enriches the reading experiences, encouraging readers to consider diverse points of view and interpretations, making expansions to human cognition of visual narratives. Through the findings, the functions of these visual narratives are found with the help of a variety of different utterances and cues that are culturally particular: these cues include different facial expressions, body language, and iconic symbols with cultural importance. Truly, the interplay among structure, grammar, and functions in visual narratives is crucial for effective storytelling.

As a result, it becomes clearer that visual narratives are not merely passive carriers of entertainment or aesthetic value; rather, they are complex systems of meaning-making that require active interpretation from the reader. While structure and grammar provide the blueprint, it is the culturally grounded functions that truly animate the visual text. These functions go beyond simply expressing emotion or exaggerating actions; they often operate as embedded codes that reflect social ideologies and historical memory.

Moreover, visual narratives offer unique possibilities that allow for sequential storytelling. The simultaneity of text and image opens up a kind of double narrative—one that is seen and another that is read—creating a space for irony, juxtaposition, or subtle cultural critique. In the case of *Hada Bhoda*, the simplicity of the artwork belies its sharp observation of Bengali familial and social dynamics. Similarly, *Doraemon's* futuristic gadgets and fantastical plots provide a medium through which issues of childhood agency, academic pressure, and urban alienation in Japan are explored. When visual grammar is employed effectively, it invites readers not just to decode the storyline but also to question the systems that underlie it.

In cross-cultural studies like this dissertation, it is also important to consider the role of the reader's background in interpreting these cues. What may seem like a trivial symbol to one audience might carry significant weight for another. A small red bindi on a woman's forehead in *Hada Bhoda* (see, *Figure 35.*) might be read as a simple adornment by a foreign reader, while for a Bengali audience, it resonates with connotations of marital status or tradition. Likewise, Doraemon's characters express emotions in ways that may appear exaggerated to those unfamiliar with Japanese manga conventions but are instantly recognisable and meaningful to native readers.

Furthermore, the medium through which these visual narratives are consumed also plays a role. Unlike digital comics or webtoons, the tactile experience of turning a page and the pause between panels invites a certain rhythm and intimacy, while scrolling through a tab changes how transitions and pacing are perceived. Understanding how structure and grammar translate across these platforms is also pretty essential to appreciating the evolving nature of visual storytelling.

This study also opens the door to considering how multilingualism interacts with visual grammar. In regions like Bengal or Japan, where dialects, regional scripts, and even hybrid languages are common, the textual components in these visual narratives are often layered with meaning. A character might speak in colloquial slang, formal speech, or code-switch between languages, reflecting socio-economic status, regional identity, or emotional state. When paired with visuals, these linguistic shifts amplify the narrative impact, offering deeper insights into character and context. Visual Linguistics provides a framework to delve into the interplay between language variety and image structure, a field that could be explored further.

Lastly, one cannot overlook the pedagogical implications of this research. Visual narratives are increasingly being integrated into educational settings, both as texts for analysis and as tools for learning. Their capacity to convey complex ideas through accessible formats makes

them ideal for engaging diverse learners. Grasping the structure, grammar, and function of visual narratives allows educators to use these texts not just for teaching language and literacy but also as powerful tools to promote intercultural understanding and to encourage critical thinking.

Therefore, the study of visual narratives through their structures, grammar, and functions, particularly with cross-cultural comparisons, sheds light on the intricate ways in which stories are constructed and interpreted. It affirms that images, much like words, are not only created with the help of their skeletal format through structures and grammar but are also carriers of meanings shaped by their cultural, historical, and social contexts. As modes of communication evolve, especially in the digital age, the need to understand this interplay becomes even more relevant. The grammar of images, their structure, and their functions are not static—they evolve with society, with technology, and with the readers' gazes.

Chapter 5

Findings (Objective 2): A Cross-Linguistic Analysis of Visual Narratives through their Temporal Aspects

5.1 Temporal Aspects in Visual Narratives

To find out how the structures of *Patua* scroll paintings and *Emaki* scroll paintings differ from the visual narrative grammar of Bengali comics and Japanese manga, respectively.

For the second objective of the study, cross-cultural temporal aspects of visual culture communication over time were explored, using data from *Hada Bhoda*, *Doraemon*, Bengali *Patua*, and Japanese *Emaki* scroll paintings, respectively. Saussure's (1916) diachronic and synchronic theories give an excellent foundation for this exploration and analysis. Diachronic research evaluates how the iconicity of achronological panels and the semantic nature of gutters have changed throughout time, whereas synchronic analysis focuses on the contemporary aspects of these visual narratives. For this, the study cross-analysed a few scroll paintings of both cultures (*Patua* scroll paintings of West Bengal and *Emaki* scroll paintings of Japan). For the *Patua* scroll paintings, I personally visited *Naya, Pingla* (a village in West Bengal, India) and collected data for the cross-analysis myself. Whereas, data on Japanese *Emaki* scroll paintings were sourced from the internet. The differences and similarities between these early scroll paintings with the modern visual narratives of *Hada Bhoda* (Bengali comics) and *Doraemon* (Japanese manga) were also identified. The study adopted a qualitative research technique, emphasising the understanding of the underlying meanings, interpretations, and cultural nuances of the data over numerical analysis or statistical measurements. In this study, primarily a content analysis method and a qualitative approach, along with a short informal ethnographic inquiry for the *Patua* scrolls, were used as mentioned in Chapter 3 for analysing the selected visual narratives, with a particular

emphasis on the diachronic and synchronic theories of Linguistics found in the bimodal form of visual narratives. By addressing the research questions, such as the ordering of panels, the semantic differences in natural gutters, and the illustration of social events over time, this study reveals how temporal factors in visual narratives impact the effectiveness and comprehension of visual communication.

Figure 41. Ranjit Chitrakar, giving me Information about Patua Scroll Paintings



I visited the village of *Naya* in *Pingla* and interviewed *Ranjit Chitrakar*, who is originally known as *Bahar Chitrakar* from *Nandigram, Purba Medinipur*. He has been living in *Naya* for more than 50 years now, since his marriage. He learnt *patachitra*, or the art of scroll paintings, as a child in poverty, and despite the adversity, he

pursued this art form that had been passed down through eight generations. *Ranjit* married into a *Patua* family, and his father-in-law, *Pulin Chitrakar* was a *Rashtrapati Award* recipient. He has perpetuated the tradition with his wife and children, who also practice the art of scroll paintings and are known as *Patuas*. The *patachitra* community in *Naya* is a recent attraction to tourists and researchers nowadays, and has gained global recognition. The oral tradition of the songs that depict the sequential flow or sometimes the random sequences of the scroll paintings is gradually diminishing, even though the new generation continues to practice the art form. Therefore, the community aspect of songwriting has decreased; it seems that *Ranjit* is the last person who still sings along while displaying the scroll paintings. These paintings are usually painted with natural and eco-friendly colours on pieces of cloth; they make these colours by gathering them from leaves, fruits, flowers, plants, trees, and other natural sources. For example, they extract saffron from *lotkon* leaves, blue from *aparijita*

flowers, brown from *segun* trees, yellow from turmeric, green from *seem* (green beans) and black and white from charcoal and *ghusum mati* (a kind of mudstone), respectively. These *Patua* scroll paintings range from an ancient variety since *Ranjit's* father-in-law's time to the modern period, i.e., now.

Similarly, as Samejima (2019) states, an *Emaki*, or the traditional Japanese illustrated scroll, is made up of two primary parts: the *kotobagaki*, which is the text that tells the story, and the drawings, which show scenes from the same story. The components of the story shift accordingly as the story progresses. Unlike the *Patua* scroll paintings that have orally sung songs along with the illustrations, the *Emaki* scrolls appear to be more similar to the modern-day visual narratives that consist of both texts and visuals. Samejima further states that the *kotobagaki* and pictures are created on thin materials like *washi* paper or silk cloth, also known as *honshi*. A backing called *uruachi* is also applied to strengthen the sensitivity of the surfaces. When the *Emaki* is rolled up, the visible cloth known as *hyoshigire* is chosen to complement the story's themes. Occasionally, a *daisen* that functions as a book cover is pasted over the surface. When an *Emaki* is opened, the very first thing that is noticed is the *mikaeshi* or the decorative paper on the back of the *hyoshigire*; the scroll is twisted around a wooden stick known as *jikugi*, which is normally of Japanese cedar and can have ends made of ivory or aromatic wood known as *koboku*; and to secure the rolled *Emaki*, a lace known as *makio* extends from the other end. According to Samejima, it requires a painstaking selection of materials to enhance and reflect the story that the *Emaki* tells, similar to the careful binding of an intricate book.

Although visuals are used both in scrolls and comics as a combination of both texts and images, these visual narratives are also quite different from one another. Yet it cannot be certainly denied that comics and manga have their roots in traditional scroll paintings. According to McCloud (1993), the origins of comics can be traced back to ancient

civilisations: he highlights how early scroll paintings in Asia, particularly Japanese and Chinese art, featured continuous narration of visuals that served as a forerunner to modern comics. Scroll paintings use the classic techniques of brush and natural colours, whereas modern comics and manga employ a wide range of visual styles, from realistic to comical. The very recent ones are also using current digital technologies. These differences illustrate the unique historical roots, cultural settings, narrative styles and creative abilities. The bimodal form of visual narratives, i.e., the visual and verbal methods used to create visual narratives like comics and manga, forms the grammar of picture stories. Panels, gutters, speech bubbles, captions, and other visuals are used together to portray a tale or narrative.

McCloud (1993) describes the lexicon of comic books as consisting of words, images and icons. He views the 'icon' as the most essential component of the comics lexicon, as already mentioned in Chapter 2. He also uses the expertise of semiotics, which has long examined signs and what they mean, to lay the foundation for the language of comics. In Linguistics, we can explore through the study of synchronic and diachronic theories that visual narratives also have their own building blocks that go beyond merely words and visuals alone; these are the basic building blocks in language. The way words are related to one another in phrases and sentences is basically how written and spoken languages convey meaning; therefore, forming the conception of syntax and semantics. McCloud categorises the changes between panels as part of the grammar of visual narratives. These changes help the readers to shift semantically from one panel to another, thereby forming the narrative in a comprehensive way.

Considering the structure and semantic shifts already shown in Figure 6. from the first objective of this research, *Patua* and *Emaki* scroll paintings, on the other hand, are usually horizontally or vertically unrolled to reveal the progression of a linear narrative. They are almost always intended to be shown and viewed in portions with the narrator (in the case of *Patua* scroll paintings) or the user (in the case of *Emaki* scroll paintings) unrolling the scrolls

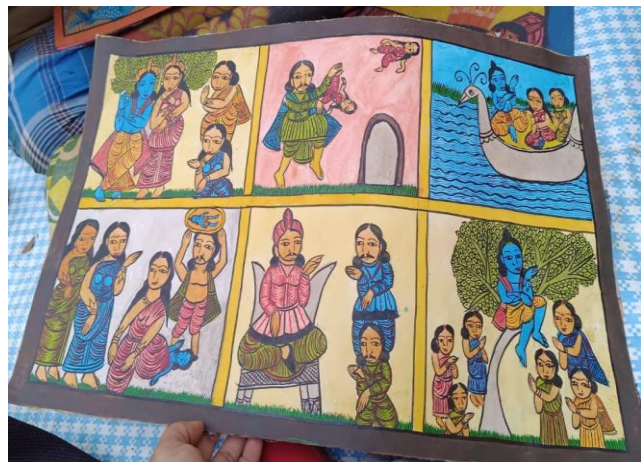
from one end to the other. The images pour onto one another almost effortlessly. There may or may not be distinct divisions of panels as defined in Bengali comics or Japanese manga. Also, each portion of the scrolls can be quite detailed, focusing on distinct aspects of a story or setting (whether narrated orally or viewed along with the texts). Rather than a series of occurrences, the emphasis is mainly on the overall composition and artistic expression.

5.2 Order of Panels in Scroll Paintings

The following are a few examples that are found in *Patua* scroll paintings, showing the order of panels (whether chronologically or achronologically) along with the natural gutters indicating contrasts with the modern visual narratives (here, Bengali comics, *Hada Bhoda*, and Japanese manga, *Doraemon*):

Figure 42., depicts an achronological order *Figure 42. Patua Scroll Painting (Achronological Ordering)*

of panels in the *Patua* scroll painting of *Krishna Katha*. The scroll displays the events with the natural time sequence, creating a non-linear storytelling experience. Nevertheless, the scroll displays the main incidents: The first row consists of—1) The love story of *Radha-Krishna* (left), 2) *King Kangsha* killing *Mahamaya* (middle), 3) *Krishna*'s pleasure trip in a boat (right);



again, the second row consists of—4) Birth of lord *Krishna* (left), 5) *Kangsha*'s informants informing about *Krishna*'s birth (middle), and the last one, 6) *Krishna* with his Girlfriends (right). The visuals, therefore, do not have any particular start, a climax, or a proper conclusion by breaking the chronological flow. While narrating the story, however, the orator points at the different panels, maintaining the chronological flow. It is also observed that there is a distinct presence of gutters between panels.

Figure 43. Patua Scroll Painting (Single Panel)

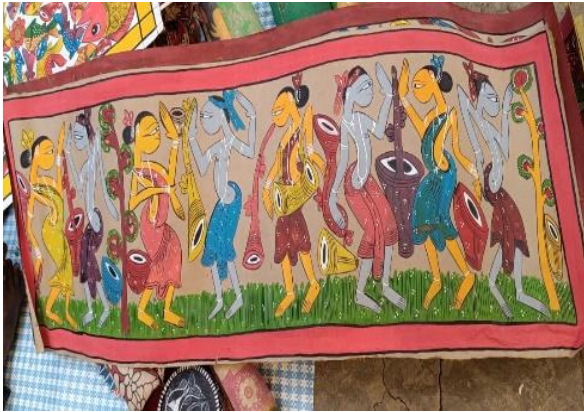


Figure 43. is a single-panel scroll painting. It is typically focused on the precise placement of a single panel within a larger storyline. The narrator narrates the whole story while pointing to the same image. The main theme of the story revolves around the event narrating about the wedding of the fish (/ma:tʰer bi:je/). Therefore, it establishes a significant moment, putting more focus and emphasis on the theme or the key information. The importance of the narrative is shifted more to the oral narration than the visual as a result of it. The visual has images of fish and some people playing instruments as a matter of celebration.

Figure 44. is a twin-panel or dual-panel scroll painting that describes two independent but interconnected frame sequences. The image and the narration are not different from Figure 4., except for the division between the two images. This has the same theme as the wedding of the fish or

Figure 44. Patua Scroll Painting (Twin Panel)



/ma:tʰer bi:je/. Even though there are two panels in the scroll painting, yet, the narrative transition or shift is not found. As a result of it, the visual focuses more on the theme than the narration.

Figure 45. Patua Scroll Painting (Chronological Ordering)

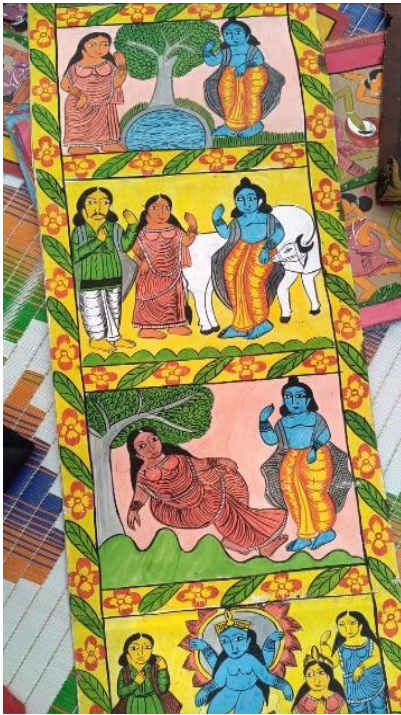


Figure 45. depicts a scroll painting with vertical chronological ordering that keeps unrolling as the narrator progresses with the story. This arrangement is similar to the vertical arrangement of panels found in the Japanese manga, *Doraemon* (see Figure 6.). It is typical of the natural flow of reading/viewing visual narratives found in many cultures, in which information is usually processed from top to bottom, making it easier for readers to follow the plot. The story is similar to Figure 42. i.e., *Krishna Katha*, unlike the achronology. This layout not only helps to retain a clear narrative structure, but it also allows for creative pacing as the vertical arrangement can be adjusted to highlight the

key points or create dramatic pauses. The gutters are well-established between the panels and also have flower and leaf patterns on them. This form helps to get a much clearer narrative framework of a scroll painting.

Figure 46. depicts a scroll painting from *Chandimangal Kavya*, or a narrative poem that has poems on different deities and goddesses such as *Chandi*, *Manasa*, *Dharma*, etc. This scroll painting uses multiple smaller panels to surround a bigger focal image that emphasises the main theme. The central image, being the main point, attracts quick attention and frequently represents pivotal scenes in the narrative. The smaller panels thereby provide context, background information, or sequential events as the narrator narrates the scenes. This structure is quite

Figure 46. Patua Scroll Painting (Multiple Panels)



effective for emphasising the significance of the primary event and highlighting contrasts with the surrounding smaller panels. This method in the *Patua* scroll paintings provides visual interest and also makes room for complex storytelling within a limited space by naturally letting the viewers move their eyes along with the story progression. The designed gutters are also well established, surrounding the big picture and also appearing between the smaller panels.

Figure 47. Patua Scroll Painting (Multiple Panels- Z-Path)



Figure 47. shows an arrangement of multiple panels like Figure 46. but in a typical Z-path arrangement similar to the Bengali comics, *Hada Bhoda* (see Figure 6.). This type of arrangement leads the viewers' eyes in a zigzag manner across the scroll. This technique entails arranging the panels in a sequential manner that begins at the top left, proceeds horizontally to the right, then follows diagonally down to the left, then continues with the flow horizontally to the right again, and so on. This structure is also quite common in comics from different other cultures (especially in the West) and can improve a narrative by deviating from traditional left-to-right, top-to-bottom reading patterns, adding visual interest, and making the story more engaging. The theme is also societal, and the art is also quite modern. Therefore, it could be inferred that this scroll painting is perhaps influenced by modern comics. The patterned gutters in between the panels are also well-established to attract viewers.

Figure 48. shows another arrangement of multiple panels with successive illustrations that flow from left-to-right, which is again very similar to the Bengali comics, *Hada Bhoda*, considering Figure 6. and Figure 7., unlike that it also goes diagonally. This visual flow allows the viewers to follow the story or the main theme that represents the scroll in a consistent and linear manner. Each section presents a vivid description of the theme of wildlife conservation that appears to the viewers as the narrator sings the narration and unrolls the scenes. It

Figure 48. Patua Scroll Painting (Multiple Panels, Left-to-Right)



is quite a modern scroll painting; the paintings in the scroll depicted in Figure 48. are half-done and yet to be completed.

Unlike *Patua* scroll paintings, *Emaki* scroll paintings follow a specific way of reading them. According to Samejima (2019), to read and view an *emaki*, it is first set on a surface level, such as the floor or a desk. The story moves from right to left, always, unlike the various types of viewing arrangements found in *Patua* scroll paintings. This typical right-to-left arrangement is also noticed in the Japanese manga, *Doraemon*, unlike it also flows diagonally (see Figure 6.). This structure of *Emaki* scroll paintings also guides its readers or viewers to navigate through the narrative in a linear manner. Reading a handscroll can almost feel quite dramatic as the viewer scrolls through a narrative from right to left, unrolling the different segments with their left hand while re-rolling the right-hand portion (Willmann, 2012). It is quite typical of scroll paintings, as it is also seen with the longer *Patua* scroll paintings. According to Willmann, the *Emaki* scroll paintings are ideal for illustrating scene-by-scene detail; they are sometimes sold in sets, allowing a large story to be extended across multiple

scrolls. Just like *Patua* scroll paintings, Japanese *Emaki* can be up to forty feet long; however, their dimensions may vary.

Figure 49. Emaki Scroll Painting- Natural Gutters and Text



Source: Wikipedia (2024)

Figure 49. is a complete picture of an *Emaki* scroll painting (*The Ippen Shōnin Eden*) with natural gutters or imaginary gutters as the images pour onto each other without a specified space in between images. It goes on as a film-like sequence as it unrolls. The text here occupies no less than two-thirds of the space; it is the seventh scroll found in the *Tokyo National Museum* (Wikipedia, 2024).

Figure 50. Emaki Scroll Painting- Alternation between Text and Painting



Source: Wikipedia (2024)

Figure 50. describes an *Emaki* scroll painting (*Hell Scroll*) with alterations between text and painting. As a result of it, gutters are made instinctively while the segments are clearly divided, giving it a panel-like outlook. This provides the necessary pauses to the readers for better reflection of the storyline. This is a 12th-century scroll painting found in the *Nara National Museum* (Wikipedia, 2024).

Figure 51. Emaki Scroll Painting- Successive Painted Scenes without Text



Source: Wikipedia (2024)

Figure 51. represents another *Emaki* scroll painting (*Chōjū-jinbutsu-giga*) with a continuous series of illustrations presented one after the other without the presence of any text. As a result of it, the story is conveyed completely through the painted scenes, creating instinctive and natural gutters. Thus, this *Emaki* scroll painting has a resemblance to the *Patua* scroll paintings, which have the absence of textual elements. It is a 12th-century scroll painting found in both the *Tokyo* and *Kyoto National Museums*, respectively (Wikipedia, 2024).

Figure 52. Emaki Scroll Painting- Successive Painted Scenes with Two Sections of Text



Source: Wikipedia (2024)

Figure 52. describes an *Emaki* scroll painting (*Heiji Monogatari Emaki*), which is also mostly a continuous flow of illustrations with only two sections of text at the beginning and the end. The beginning seems to have more written elements than the ending. Like Figure 50., the painted scenes are conveyed through the instinctive, natural gutters. It is a 13th-century scroll painting found in the *Tokyo National Museum*, *Seikadō Bunko Art Museum*, and the *Museum of Fine Arts, Boston* (Wikipedia, 2024).

5.3 Social Contexts

This study delves into how the structures of these scroll paintings also set the patterns for reflecting and shaping social settings. The social settings of scroll paintings have changed over time. Modern visual narratives (comics and manga) and also modern *Patua* scroll paintings have different social contexts from those that were painted previously. This study has analysed a few examples to look into the similarities and differences found in these visual narratives. The structure and grammar of the narratives shown above are more than just language events; they influence and reflect culture and identities also through social contexts, as already inferred in Chapter 4 through its findings.

The *Patua* scroll paintings include a variety of themes (old and new). These themes frequently draw on *Hindu* mythologies like the *Ramayana*, the *Mahabharata*, *Krishna Katha*, *Chandimangal Kavya*, and so on. Some are for entertaining purposes, and some also have social meanings:

Figure 53. The Ramayana- Old Patua Scroll Painting (Left); New Patua Scroll Painting (Right)



Figure 53., shows the two *Patua* scroll paintings of the same theme by unrolling the story of the *Ramayana* from two different periods. The picture on the left is believed to be one of the oldest *Patua* scroll paintings, probably from the 13th or 14th century, that has still been preserved. Whereas, the picture on the right is a newer one painted by *Ranjit Chitrakar*, maintaining the *Patua* tradition of narrating mythologies.

Figure 54. *Debnath's (1962) Hada Bhoda Samagra (p. 423)*



The *Patuas* traditionally wandered from place to place, narrating mythological stories through their scroll paintings to educate the people. Illiterate villagers could learn about these mythologies through their art. However, in contemporary society, these mythological stories are staged in style as depicted in *Hada Bhoda* and are enjoyed as entertainment across various art forms; The panel in

Figure 54. shows *Bhoda* weeping vigorously, seeing the tragic scene where *Rama* laments over *Laxmana's* death. It is assumed that people are already familiar with the stories, so only parts of the scenes from the *Ramayana* are presented in the comics. Consequently, the popularity of *Patua* scroll paintings has declined as people seek modern and entertaining content.

Figure 55. describes a very common theme that is observed in the *Patua* culture. There were quite a number of scrolls that represented the wedding of the fish or /ma:tʃʰer bi:je/ as mentioned earlier (see Figure 43. and Figure 44.). These fish stories generally are found in one or two panels, but the story narrates a vivid picture that makes the viewers roll their eyes all over the scroll. The narrator sings the story of the anticipated wedding of the *dariya* fish, where all the other fish were also invited. Amidst all this merrymaking, *boal*, the monster fish,

Figure 55. *Patua Scroll Painting (Machher Biye)*



watched the festivities from afar. As this fish is known for its huge size and big appetite, it was not invited. But suddenly, it decided to get into the celebration and began swallowing all the other fish that were invited, and as a result, panic started spreading through the crowd. It echoes Hobbes' (1968) 'Leviathan', where he explains how a big fish eats all the small fish. Here, in the *Patua* scroll painting, the *boal* fish represents the leviathan eating all the small fishes of different species. The theory in 'Leviathan' was not known to the *Patuas*, yet this thought matches with that of the traditional *Patua* narration.

Figure 56. Debnath's (1962) *Hada Bhoda Samagra* (p. 214)



It is also to be noted that fish have always played a big role in the Bengali culture. Figure 56. shows that a ghost is stealthily eating quite a big fish that *Hada* and *Bhoda* caught. Bengali literature and culture are heavily reliant on ghosts. As already mentioned in Chapter 4, Figure 35., there are quite a few kinds of ghosts found in Bengali literature, among them, *mechho bhoot* are the ghosts that are fascinated with fish and can 'beg, borrow, steal or

kill' for a fresh slice of fish, especially if it is hilsa (Maiti, 2015). Moreover, Bengali people are very fond of fish, and fish is not only the most commonly consumed protein but is also the preferred one (Akter, et al, 2019). So, whether it is the age-old *Patua* scroll paintings or the modern Bengali comics, there has to be some reference to fish, as it is almost part and parcel of the Bengali culture.

Figure 57. Patua Scroll Painting (Marriage and Society)



Figure 57. shows *Ranjit Chitrakar* holding a *Patua* scroll painting in different stages of its narration that depicts the story of society as it was in the past. This scroll painting, created by his father-in-law, *Pulin Chitrakar*, represents the behaviours and norms of the time. *Ranjit* is embarrassed as he describes the customs and attitudes of men and women in that era: previously, the groom, even with a squinted eye, was considered handsome simply because he was male; meanwhile, the bride, a mere twelve-year-old girl, was forced to marry a man twice her age because early marriage was a compulsion for Bengali girls. In his song narration, *Ranjit* laments the fact of modern times (during *Pulin's* era) that even at the age of sixty, a lady still follows traditional *Hindu* Bengali rituals and customs while remarrying. He also mourns that parents are neglected by their children, which has devastating implications, such as the father committing suicide. He also emphasises societal changes, such as the government's policy for restricting couples to two children, and criticises how a wife now frequently insists on going to the movies with her husband while neglecting her household duties.

Figure 58. Debnath's (1962) *Hada Bhoda Samagra* (p. 308)



Although the social context differs nowadays, girls no longer get married at an early age or get criticised when they behave modernly. It is evident in Figure 58, where girls can fight with boys and are equally mischievous and playful. Furthermore, the author, Debnath (1962), does not give the readers a negative impression of girls in his comics; in fact, it shows their rebellious and courageous nature. Still, *Pulin's* paintings are considered authentic even though the *Patuas* used to preach and narrate the evils of society through their visual narratives as they reflected their times.

Figure 59. *Genji Monogatari Emaki- Yadorigi Gi Chapter and Text Calligraphy*



Source: Wikipedia (2024)

'The Tale of *Genji*' (Figure 59.) is regarded as the world's first novel and Japan's most renowned literary work. It was written in the early 11th century by *Murasaki Shikibu*, who

was an imperial court lady. The scroll spans 54 chapters and describes events over 70 years that surround the life of *Genji*, the shining prince. It is a romance story that has over 430 people embroiled in love, envy, and ambition, all set against the backdrop of the rigorous and traditional Japanese court (Wikipedia, 2021).

Over time, Japanese visual narratives have taken on different themes. In Figure 60., a small scene from the Japanese manga, *Doraemon*, depicts a fun prediction of *Nobita's* future marriage that gives rise to different emotions among the characters altogether in these few panels. As we see that 'The Tale of *Genji*' digs into profound human emotions and *Heian-period* social systems, *Doraemon*, on the other hand, also deals with a wide range of different emotions from characters, but focuses more on futuristic gadgets and moral lessons. Nevertheless, the characters are equally engaged in the emotions of love, envy, and ambition, but in a playful manner.

Figure 60. Fujio's (1974) *Doraemon- Volume 1 (p. 13)*



Figure 61. Emaki scroll painting- Chōjū-giga



Source: Wikipedia (2024)

The *Chōjū-jinbutsu-giga* (Figure 61.) or ‘Animal-person caricatures’, is a rare collection of four scrolls from the 12th and 13th century Japan, as mentioned earlier in Figure 51. These scrolls are excellent examples of the Fujiwara period. They show frogs, rabbits, and monkeys dancing, swimming, and communicating like humans. Some regard *Toba Sōjō*'s work, with its simple black-and-white art technique, to be the first example of manga and Japanese animated films (Maruzen-Yushodo Co. Ltd., 2004). It can be noted that with this theme, the scrolls represented the literary device, ‘personification’. By using this technique, the whole narration came out to be more lively and impactful, even though *Chōjū giga* scrolls lack textual elements. The animals are used humourously to criticise and satirise human behaviour and social norms.

Figure 62. Fujio's (1974) *Doraemon*- Volume 1 (p. 27)



In the *Doraemon* manga series, the chapter ‘*Henshin Crackers*’ (Figure 62.) presents edible crackers that allow the characters to morph themselves into animals. Here, we observe a reverse factor of what we notice in *Chōjū giga*

scrolls; humans turn into animals. This reflects the distinct artistic and narrative traditions of Japanese visual narratives across time.

Therefore, the structures of panels and social contexts present in the visual narratives are again found to be closely intertwined, even when examined through the diachronic aspects of picture stories. This notion provides much stronger support for the findings in Chapter 4. As visual storytelling evolves over time, the structure often reflects shifts in social contexts, revealing how societies adapt and change.

5.4 Discussion

The ability to comprehend visual narratives is much more than just a cultural nuance; it is a universal cognitive function that serves as the very foundation for human communication, understanding, and social interaction. By expanding this study to the Saussurian concepts of diachronic and synchronic factors in Linguistics, it is well observed through the findings how visual narratives have evolved and function in a very similar manner to languages. Derrida argues that history has changed our comprehension of structure by showing that the central points are not fixed and that they can shift over time. Events, changes, and new ideas challenge these centres, causing ruptures in how we perceive and define structures (Almabrouk, 2020). Therefore, history reveals that structures are not absolute or static. They keep evolving and are thus open to reinterpretations. By relying on Derrida's belief, it is understood that we can move beyond limiting ideas of structure and embrace different variations and openness in how we understand and interpret systems. The evolution of visual narratives, from *Patua* scroll paintings to Bengali comics or from *Emaki* scroll paintings to Japanese manga, corresponds to the diachronic development of language. Just as languages acquire complex frameworks to convey meaning, visual narratives, too, have evolved their sequential approaches through their panel formations and story progressions to improve comprehension and engagement. Such sequential narratives demonstrate continuity and change across time. The transition from scroll paintings to modern visual narratives or

modern scroll paintings (in case of *Patua* scroll paintings) that have inherited the usual comics' panels, shows developments and adjustments in narrative approaches. Furthermore, visual narratives can also be analysed synchronously at any time for a better understanding of their structure and functionality. Different cultures represent distinct yet universally understandable sequential frameworks in their visual narratives. Synchronic analysis reveals these differences as well as their similarities, much like how different languages have unique but fundamentally similar elements. For example, communication among animals like humans is a very common thing in almost all cultures. Like we see in Figure 55., Figure 61. or Figure 62., the fascination with animals behaving as humans is quite prevalent in different cultures, dating back to Hobbes' (1968) times with his Leviathan theory. Several scenes from Egyptian tombs also depict animals as actors performing human roles, behaviours, or occupations that were recreated on ostraca or postcards (Ezzat, 2021). The study of visual narratives, as a result, provides a unique lens through which we may examine certain transitions as well as consistencies, demonstrating how narrative structures, along with social situations, evolve and remain the same in some cases across time.

Through diachronic linguistics, visual narratives reflect a progression of artistic and narrative traditions influenced by historical and cultural contexts. For instance, the shift from mythological narratives of *Patua* scrolls to the staged style depicted in *Hada Bhoda* or the satirical themes of *Chōjū-giga* scrolls to the modern entertainment of *Doraemon*, emphasises the temporal transformation of visual language through reader preferences.

Likewise, synchronically, visual narratives provide a snapshot of cultural and social structures at a given moment: the *Patua* scroll paintings reflect the socio-religious fabric of their time, where art served as a tool for education and oral tradition. Similarly, *Chōjū-giga* scrolls, with their playful personification and critique of societal norms, mirror the values and humor of the *Fujiwara* period. Modern Bengali comics and Japanese manga capture the

aspirations and fantasies of contemporary audiences while reinterpreting traditional themes in innovative ways.

Narratives, whether in the form of myths, legends, epics, or contemporary comics or manga, are not merely aesthetic or entertaining artifacts. In fact, they are evolving the semiotic systems that encode cultural, social, and historical knowledge. Turner (1986) critiques the rigid formalism of structuralist approaches, arguing that narratives are not just fixed sets of structural oppositions but also dynamic meaning-making processes, shaped by context, performance, and audience reception. By applying Saussurean linguistic principles, this study can build upon Turner's criticisms, demonstrating that visual narratives, just like language, undergo diachronic shifts and reflect synchronically embedded cultural frameworks. Therefore, it reveals how visual storytelling is not static but constantly negotiating traditions and contemporary realities. It also shifts the focus from creator-centric structuralism to audience-driven evolution, highlighting how readers actively reshape narrative structures. From scroll paintings to contemporary visual narratives, picture stories function as a living system of signs, constantly redefining their referents through time.

But it should also be kept in mind through the study of Saussure's (1916) panchronic linguistics that, despite cultural and historical frameworks, scroll paintings and modern visual narratives follow sequences of images typically for the narrative flow. Whether it is a *Patua* scroll or a Japanese manga, the fundamental principle of how a series of images tells a story remains consistent across times and cultures, leading us to a broader understanding of universal principles of how human cognition processes visual information and narrative structures.

This temporal aspect of diachronic and synchronic perspectives in visual narratives offers a compelling framework for examining how stories evolve and endure. While historical transitions in visual form and content are clear—from religious instruction to social satire and

entertainment—what's equally fascinating is the continuity in the cognitive mechanisms used to decode these stories. Even when aesthetics shift or societal values change, the visual logic behind sequencing, symbolism, and character representation often remains surprisingly familiar. This suggests that while cultural expressions are bound by their time, the human impulse to narrate visually, to sequence images in a meaningful way, cuts across generations.

Moreover, audience reception plays a crucial role in shaping visual narrative traditions. In many cases, the popularity of certain formats or styles isn't solely about their artistic merit but their accessibility and relevance. For example, the humour in *Hada Bhoda* appeals to everyday sensibilities in West Bengal, while *Doraemon* reflects the anxieties and hopes of a hyper-technologised Japanese childhood. This responsiveness to the audience suggests that visual narratives are not just products of their time, but are also shaped together with their readers. The narrative grammar adapts not just in form but in tone, pace, and even moral perspective. A *Patua* scroll might have once guided its viewers through religious parables; today's comic panels may instead pose ethical dilemmas grounded in modern life—both, however, rely on culturally encoded signs to do so.

In this light, the evolution of visual storytelling becomes less about linear progress and more about a flexible system that expands, contracts, and reshapes itself in dialogue with its cultural context. Scroll painters and comic or manga artists, separated by centuries and continents, may unconsciously employ similar semiotic strategies, such as exaggerated gestures or emblematic objects, because these are embedded within collective human visual consciousness. This could be why even young readers, unfamiliar with the historical backdrop of a scroll painting or a vintage comic strip, can still 'read' them and derive meaning.

To summarise, this adaptability points to the resilience of visual narratives in the digital age. Just as oral traditions once transitioned into scrolls and then to printed comics, today's

memes, GIFs, and interactive webtoons are a continuation of the same impulse—to distill complex ideas, emotions, or critiques into accessible visual forms. Despite the medium's transformation, the cognitive process of interpreting signs, understanding a facial expression, deciphering an action sequence, or catching a visual pun remains rooted in age-old practices. Thus, studying visual narratives across time not only uncovers shifting artistic conventions but also deepens our understanding of how humans have always made sense of the world through picture stories.

Chapter 6

Findings (Objective 3): A Cross-Linguistic Analysis of Visual Narratives through Speech Act Theory

6.1 Speech Acts in Visual Narratives

To explore how visual culture is communicated in the visual narratives of *Hada Bhoda* and *Doraemon* through different communicative acts, facial expressions, ideophones or onomatopoeic words and drawing styles.

Austin's (1959) speech act theory provides a useful framework for what this research explored through the third objective of the study. The visual culture is communicated in the visual narratives of *Hada Bhoda* and *Doraemon* via facial expressions, ideophones and drawing styles. According to this notion, visual features can communicate in ways that are analogous to verbal discourse. As a result, this research is concerned with understanding how facial expressions convey emotions and cultural nuances, how onomatopoeic words or ideophones change in each narrative, and how multimodality plays a role through the drawing styles. By exploring these elements with the help of a content analysis method and a qualitative approach, this study uncovered how visual and textual cues act as performative expressions of culturally specific meanings in visual narratives.

6.2 Communicative Acts

The speech acts in Austin's (1959) speech act theory are understood as utterances through which a speaker communicates and conveys a specific meaning that the listener is meant to understand. Speech acts are mainly classified into three main types: firstly, locutionary acts are the acts of saying something, basically the acts of producing the utterances that carry meaning. Secondly, illocutionary acts go a step further, as they convey the speaker's intentions. For example, when a speaker makes a request or commands, or promises

something, they are performing an illocutionary act. Lastly, perlocutionary acts are the acts that focus on the impact or effect that is made on the listener by the utterances.

Concerning the illocutionary acts, Searle (1976) expanded on Austin's (1959) speech act theory by developing a refined taxonomy of speech acts, organised into five main categories based on distinct functional dimensions. They are as follows:

- i. Assertives:** The speaker commits to the truth of what is asserted, conveying information or beliefs that include statements, claims, descriptions and predictions.
- ii. Directives:** The speaker makes an attempt to get the subject to do something by expressing their wish, including orders, requests and instructions.
- iii. Comissives:** The speaker commits to take action in the future with promises, oaths and bets.
- iv. Expressives:** The speaker expresses a variety of psychological states that include greetings, congratulations, apologies, awe, etc.
- v. Declarations:** The speaker brings about an immediate change in a situation with hiring decisions or war-like declarations.

This taxonomy by Searle (1976) is often seen as an improvement on Austin's (1959) speech act theory, as it avoids overlaps by grounding classifications in a consistent set of dimensions. Furthermore, Searle emphasised that speech acts are embedded in a 'network' of unspoken beliefs and a 'background' of capacities, thereby making this context critical for the speech act's meaning to be fully realised. For instance, the promise, 'I will buy you a drink' relies on shared understandings about money, dining and social conduct.

Application of this taxonomy into visual narratives involves the interpretation of visual elements as forms of communicative acts that convey meaning beyond just spoken or written languages. A thorough analysis of Searle's (1976) taxonomy within the Bengali comics, *Hada Bhoda* and the Japanese manga, *Doraemon* has been conducted through a cross-

linguistic analysis. The following analysis explores how the different classifications, like assertives, directives, commissives, expressives and declarations, reveal the uniqueness of the culture-specific visual features present in the visual narratives that convey meaning and engage with the readers through distinct ‘communicative acts’:

Figure 63. from the *Hada Bhoda* series observes *Hada* noticing a spider as he shouts out expressing his fear, ‘Oh Gosh! A Spider!’ This utterance communicates shock or mild fear, which is evident in the speaker’s facial expression and

Figure 63. Debnath’s (1974) *Hada Bhoda Samagra* (p. 252)



emotional reaction upon encountering a spider. The phrase, ‘Oh Gosh!’ adds to the statement and serves to convey *Hada*’s negative psychological evaluation of the spider, that it is creepy. In the immediate next panel, it is observed that *Bhoda* is making a claim that also reflects on his body language based on what he has observed about *Hada*: ‘Now it is clearly understood what *Hada* is scared of.’ It is an evidence-based assertion stating that *Hada* is afraid of spiders and conveys information that reflects an assumption about *Hada*’s characteristic trait, presenting it as a fact. Together, these two panels create an **expressive** utterance, sharing *Hada*’s subjective feelings and an **assertive** utterance that represents an assumption that *Bhoda* believes to be accurate.

Figure 64. Debnath's (1962) *Hada Bhoda Samagra* (p. 328)



In Figure 64., it is observed that *Hada* is awestruck at the amount of medicine that he has been served by his uncle. His uncle says, ‘It would do no good if not taken the whole glass. Now please hold the glass.’, which aligns with the purpose of a directive speech act. It gives a reason for the directive by encouraging the listener to follow through by presenting a condition that the medicine would not work if not taken the entire glass. Furthermore, the

uncle puts more stress on holding the glass immediately, which also directs *Hada*, ensuring the instruction is followed specifically, as if like a command or an order with a serious facial expression. Therefore, the uncle’s utterance highlights his desire for a particular response from *Hada*, explicitly guiding him to do so, making the sentence **directive**.

In Figure 65. *Hada* and *Bhoda*’s friend commits to a future action where he says, ‘Don’t worry! I will treat you guys here. What do you guys want to eat?’ It is a clear commitment by the speaker to cover the cost of the meal, and as a result, it serves as a binding statement that indicates the speaker’s intention to follow through. *Hada* and *Bhoda*’s friend asking them about their choices of food and

Figure 65. Debnath's (1962) *Hada Bhoda Samagra* (p. 329)



Hada and *Bhoda* replying with Bengali delicacies like *chop cutlet* (snacks) and *doi sondesh* (sweets), emphasises the speaker’s dedication in the statement to fulfilling the promise, which is also observed through his bodily gestures. Therefore, it makes a **commissive** statement by focusing on the speaker’s willingness to perform his action by treating his friends.

Figure 66. Debnath's (1962) *Hada Bhoda Samagra* (p. 346)



Figure 66. depicts a classroom scenario where the teacher says, 'It is a difficult test. The principal has decided that the boys who would pass this test would be given prizes.' This utterance is a formal declaration in the classroom setting, where the teacher talks about the effects on the students based on the principal's authority, indicating that this statement has the power to enact a new reality regarding the reward policy. It is also very evident from the body language of the teacher that he is pointing his finger towards the students. Therefore, it directly establishes a new condition or sets a rule, thereby making the sentence a **declaration**.

Figure 67. depicts an outdoor sports scene from the *Doraemon* series where *Nobita's* father is providing direct instruction to the listener (here, *Nobita*) on how he should position himself, specifying what he needs to adjust; he directs

Figure 67. Fujio's (1974) *Doraemon- Volume 5* (p. 49)



him regarding his body posture and also tells him how he should execute the throw of the ball, emphasising the desired action concisely and directly. The father also has a commanding attitude through his facial expressions. In response, *Nobita* asserts that the weather does not seem favourable to him, and it is cold as the temperature is low. As a result, this utterance presents an observation about the weather, which *Nobita* perceives as factual information. Providing an extra phrase by adding that his hands have become numb, the speaker connects the cold weather to his physical reaction, grounding the statement more in a logical and factual format. His facial expressions and bodily gestures further describe his unwillingness

to practice the sport because of the cold. Therefore, the panel gives two examples of **directive** and **assertive** utterances, providing clear instructions to achieve the intended result, and representing the situation as objectively as possible, respectively.

Figure 68. Fujio's (1974) *Doraemon- Volume 5* (p. 48)



Figure 68. depicts a scene where *Nobita* is seen happily walking by while a speech bubble from his father (who is not present in this particular panel) puts a statement that *Nobita* will be treated as he was treated by his father. Here, the speaker is making a commitment to act in a particular manner. It sounds like a promising signal of an obligation that the father is intentionally taking on. The added phrase

'the very way my father treated me' clarifies the nature of the intended treatment, showing that the speaker is committing to a specific standard that he has in mind. This typical approach of putting up the utterance, even though the facial expressions are not visible, makes the utterance a **commissive** sentence, clearly establishing a commitment from the speaker, thereby binding himself to follow through on his words.

Figure 65. Fujio's (1974) *Doraemon- Volume 2* (p. 77)



In Figure 69., *Doraemon* seems to be yelling at *Nobita* for leaving something at *Gian*'s house, whereas *Nobita*'s facial expressions with raised eyebrows and a sorrowful face do justice to his utterance. He expresses a personal reaction of empathy towards his friend, *Gian*, which expresses his emotional response when *Doraemon* seems angry. It also brings out his psychological state of

feeling sorry and expresses it through the depth of his concern. Therefore, the utterance qualifies as an **expressive** statement as it communicates the speaker's feelings about Gian's situation, allowing Doraemon to understand his empathetic reaction.

Figure 66. Fujio's (1974) *Doraemon- Volume 4* (p. 103)



Figure 70. observes an authoritative position of *Gian*, where he declares the power to bring about a real change or consequence for disobedience. His mischievous smile makes him look fascistic as he establishes a new rule that becomes effective through his utterance, which stresses the fact that

anyone who disobeys him will face execution. Therefore, the utterance clearly qualifies as a statement of **declaration** as it institutes a new condition, enforcing consequences on those who disobey the speaker, thereby changing the social reality based on his words alone.

Speech bubbles represent the locutionary acts in visual narratives, while illocutionary acts are reflected in how the text within these bubbles, combined with facial expressions, gestures, and body language, conveys intentions and attitudes. Therefore, a complete interplay of visual and textual elements works together to communicate meaning beyond the words themselves. On the other hand, perlocutionary acts reflect the reactions of the characters within the narrative as well as the responses elicited from readers, shaping the emotional and cognitive engagement with the visual narratives. This process emphasises multimodality in communicative acts, as visual narratives depend on multiple modes like texts, images, and other symbols to enhance overall comprehension for the readers. Through the integration of visual cues like facial expressions and gestures, multimodality enriches the communicative potential of the visual narratives. With the cross-linguistic lens of the Bengali comics, *Hada*

Bhoda, and the Japanese manga, *Doraemon*, multimodality demonstrates how different modes of communication are combined to resonate with culturally diverse readers, making it a rich field of research by exploring cultural nuances and shared human experiences across the domain of visual storytelling.

6.3 Facial Expressions

Facial expressions are considered to be a universal form of indirect or nonverbal communication. Although their interpretation varies greatly across cultures, comprehending these differences is critical for effectively detecting emotional and social cues in a multicultural atmosphere. Darwin (1872) argued that facial expressions of certain basic emotions, for example, happiness, sadness, fear, wrath, surprise, and disgust, are biologically hardwired and therefore are universal across cultures. He hypothesised that these facial expressions have evolutionary roots and have adaptive purposes that help in survival and social communication. For instance, wide eyes of astonishment boost visual perception, and exposing teeth in anger might be used as a threat display. According to Darwin, these expressions are innate and are passed down through generations, representing the history of human evolution. Later, Ekman (1971) considerably contributed to this theory of universal facial signs of emotion. His study emphasises the interaction of biological universals and cultural specifics in the expression and interpretation of emotions. The universal features thus serve as a shared base for emotional communication, while cultural differences add layers of complexity, influencing how emotions are expressed, perceived, and controlled in different social situations.

Despite the universality of these emotions, their intensity, frequency, and social appropriateness vary from culture to culture. These expressions are therefore influenced greatly by the way they are expressed and interpreted. To effectively analyse how visual culture is conveyed through facial expressions in *Hada Bhoda* and *Doraemon*, the basic

human emotions are examined in this study and found to be present in these two visual narratives. Visual culture refers to how visual pictures convey cultural and social messages. Hence, facial expressions play a significant role as a part of reading visual narratives and understanding the culture through them. These expressions are a powerful strategy for transmitting visual culture. They not only represent the characters' mental states but also portray all the inbuilt cultural and societal issues. Examining certain examples from *Hada Bhoda* and *Doraemon* reveals how the two artists have employed facial expressions to communicate complex emotional and cultural messages in order to progress the storyline.

The major facial expressions conveyed through *Hada Bhoda* and *Doraemon* are follows:

Figure 67. The Facial Expressions found in Debnath's (1962) Hada Bhoda Samagra

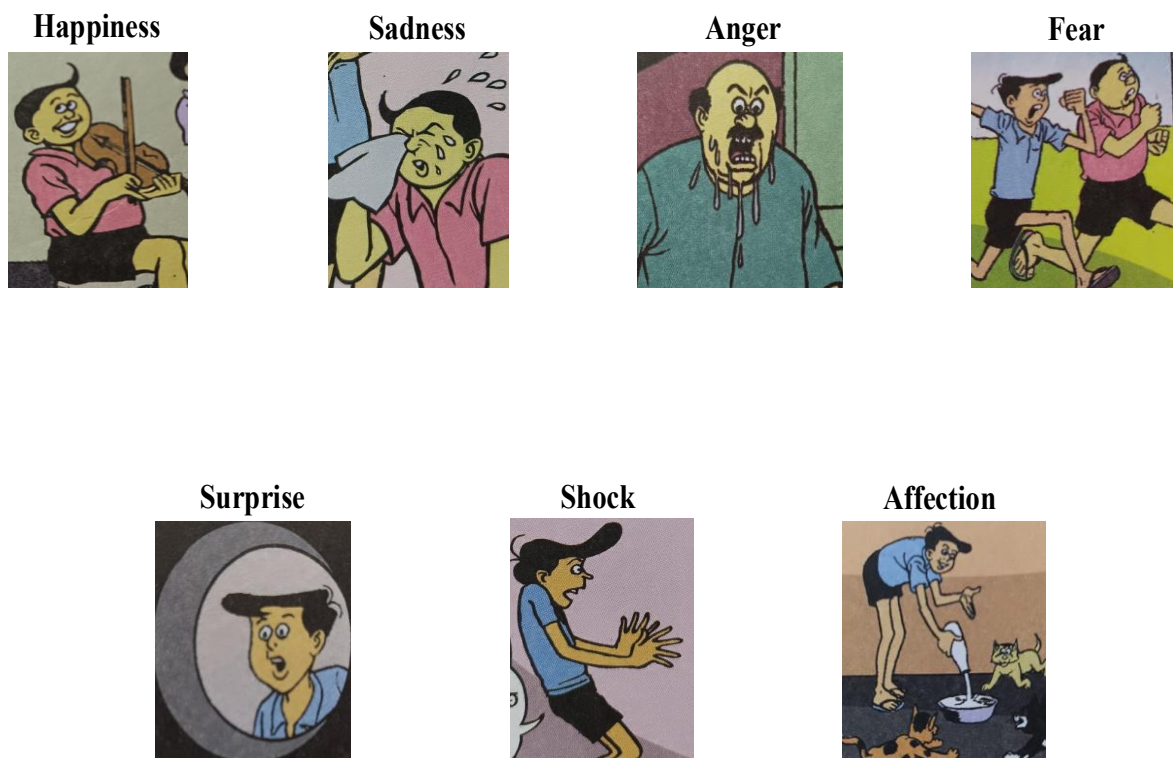
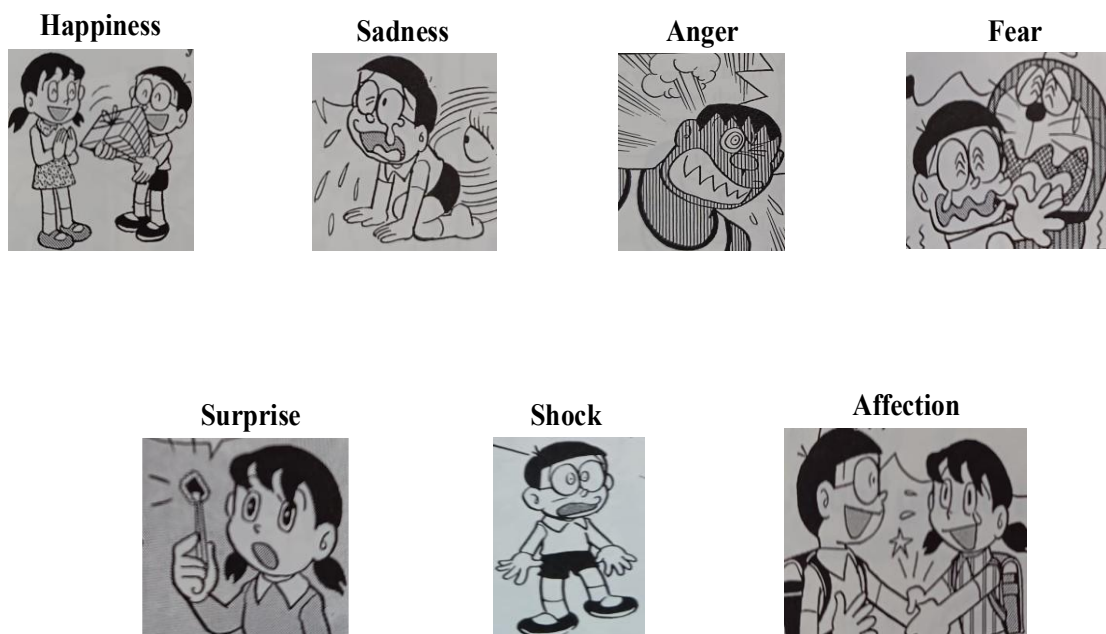


Figure 68. The Facial Expressions found in Fujio's (1974) Doraemon series



Both the visual narratives from the two different cultures of Bengal and Japan have a similar range of facial expressions. The expressions could be interpreted easily by people across cultures, yet there are a few differences in the drawing styles to convey the visual cultures:

Happiness: It is portrayed through the cues that convey a sense of joy. It is found that both in *Hada Bhoda* and *Doraemon*, the faces are smiling and the eyebrows are raised. *Bhoda* shows a more relaxed posture with bright and joyful open eyes, whereas *Shizuka* and *Nobita* look like happy individuals with fluid movements.

Sadness: It is portrayed through the different visual elements that express a sense of melancholy. In the above images, it is found that both in *Hada Bhoda* and *Doraemon*, *Bhoda* and *Nobita* have tearful eyes, respectively. Weeping *Bhoda* is observed to have been wiping off his tears with a handkerchief while his eyes are closed, whereas *Nobita* has extremely watery eyes with one eye closed. He seems to be crying loudly enough as his mouth is open wide.

Anger: Anger is often associated with aggression and frustration. Similarly, in *Hada Bhoda*, *Hada* and *Bhoda*'s uncle seems to be quite frustrated with frowning eyebrows, intense eyes, and aggressive gestures with an open mouth and teeth showing within. On the other hand, *Gian* from *Doraemon* has similar frowning eyebrows, but one of his eyes is closed, and the other is quite big and rolled over, portraying a more intense form of wrath. He also has a flushed face (shown with shades of vertically inked lines), a wide mouth and baring teeth. The visual trope or the multisensory effect of the smoke that is coming out of his head adds more to the fierceness of his anger.

Fear: Fear generally portrays a sense of terror and unease. In *Hada Bhoda*, *Hada* and *Bhoda* are observed to be running hastily out of fear with wide eyes, raised eyebrows, and an open mouth, signifying panting. On the other hand, in *Doraemon*, *Doraemon* and *Nobita* are observed to be too afraid as they almost have teary eyes. Their curly mouth and closed eyes layered three times, one after the other, intensify the tension and the moment of fear that they are facing as they also hug each other tightly.

Surprise: Surprise typically represents an unexpected or sudden positive reaction of awe. Not many differences are found in the visual cues presented by *Hada* and *Nobita* in the above pictures from *Hada Bhoda* and *Doraemon*, respectively. They both have raised eyebrows, wide eyes, and open mouths, depicting awe.

Shock: Similar to surprise, shock, too, is an unexpected reaction to some emotional disturbance. Just like surprise, shock also has similar facial expressions as found in *Hada Bhoda* and *Doraemon*, where *Hada* and *Nobita* both have wide eyes, raised eyebrows and open mouths. What makes it different from surprise is the body posture. *Hada* has a backward movement in his posture with his hands towards him, saving himself from some attack, as if physically he is flinching from the shock. Whereas, *Nobita* has a stiff or momentary frozen posture as if he is unable to process some information.

Affection: Affection depicts love and emotional connection between individuals. In *Hada Bhoda*, *Hada* has a smiling face with raised eyebrows and a soft gaze as he feeds milk to the kittens. His body is also leaning towards them, showing emotional engagement. In *Doraemon*, too, *Nobita* and *Shizuka* have smiling faces, raised eyebrows, and soft gazes towards each other. *Shizuka*'s affection towards *Nobita* is more enhanced with tears coming out of her eyes, and it shows closeness as she holds his hand with care.

In visual narratives, such emotions as happiness, sadness, anger, fear, surprise, shock, and affection are significant features of communicative acts, adding more intense flow to the story progression. These expressions in characters are not just reactions but important intentional and visual cues that convey their thoughts, intentions, and relationships to that of the readers, molding how the narrative unfolds. These emotional expressions link with other communicative elements in the stories, such as speech bubbles, gestures, and body language, providing a holistic communicative factor and letting the readers go deep into the characters' inner worlds and intentions. Therefore, these expressions are crucial in visual narratives as they help build up the stories, allowing them to progress in a meaningful manner. The expressions also keep the readers engaged and entertained throughout each character's journey.

6.4 Ideophones

Facial expressions in visual narratives are significant components of nonverbal communication because they smoothly convey a wide range of emotions that help the readers comprehend the story progression. These visual signals are often recognisable and rapidly transmit the characters' emotional states. Similarly, onomatopoeic words or ideophones help to bring visual narratives to life. These words phonetically replicate or mimic natural sounds in human communication. Just as facial expressions physically indicate emotions, ideophones audibly represent actions and sounds through visual cues, enhancing the readers' sensory experiences. Unlike conventional words, ideophones or onomatopoeias frequently convey

rich imageries and elicit a visceral or emotional reaction in listeners. This study has observed some onomatopoeic sounds from the visual narratives of *Hada Bhoda* and *Doraemon*, and also noticed a few differences with similar pictures:

Here in Figure 73., *Bhoda* and his friends are mocking *Hada* by comparing him to frogs. He exclaims, ‘(He is) playing frog-jump with real frogs.’, which simply describes the activity of playing, involving frogs. Here, the Bengali ideophone is more significant than *Bhoda*’s utterance, which vividly captures the sound and sensation of frogs: /krojã:k kôja:k/. The sound /krojã:k kôja:k/ could be influenced by the most common noise, ‘croaking’ in English, that a frog makes.

Figure 69. Debnath’s (1962) *Hada Bhoda Samagra* (p. 38)



Figure 70. Debnath’s (1962) *Hada Bhoda Samagra* (p. 415)



In Figure 74., *Hada* is trying to aim at the crow with his shooting gun. But his shot missed the target, and he exclaims with disappointment that it hit the bird’s nest instead. As the bird takes off its flight, it is accompanied by a sharp, piercing ‘cawing’ sound, /ka: ka:/, which is a common representation of a crow’s call in Bengali culture. It typically imitates the

sound of a crow and is quite similar to the English onomatopoeia, ‘caw’.

Figure 71. Debnath's (1962) *Hada Bhoda Samagra* (p. 378)

In Figure 75., *Hada's* father is seen to be up in the air with *Hada's* mixed expressions of shock and fear. It is also noticed here that some springs have erupted from beneath the fabric of the bed mattress. As a result of it, a metallic sound is used to imitate the sound of the springs that made *Hada's* father bounce: /**poɔiŋ**/. This sound is also quite similar to the ideophone, 'boing', commonly used in English



comics when depicting a bouncing object or character. In both cases, bilabial plosives /p/ and /b/ are used. The Bengali onomatopoeia is voiceless (softer), with English being the voiced (harder) one.

Figure 72. Debnath's (1962) *Hada Bhoda Samagra* (p. 68)



Here, in Figure 76., *Bhoda* seems to be taking an act of revenge on *Hada*, as is evident from his dialogue where he is saying that it is a 'tit for tat' as he could push him well. On the other hand, *Hada* seems too confused as he is almost falling off his playing cart. While they are in this playful act, a distinctive ideophone is noticed: the

collision of two wooden carts produces a very common onomatopoeic sound in Bengali: /**gʰɔtaŋ**/. This is typically the 'clacking' sound, which happens because of the involvement of two robust and heavy wooden objects.

Figure 77. captures a sudden and messy moment when a bowl full of thick yogurt slips off the hands of the man in green and blue, and hits the face of the other man. This is an accidental incident caused by the kite. The thick and creamy yogurt hitting the man's face and spreading all over makes an onomatopoeic word in Bengali, /sljæpa:t/. This sound resembles the English ideophone, 'splat', which similarly describes the sound of something wet and soft hitting some surface; it occurs that the /sl/ and /sp/ sounds are interchanged.

Figure 73. Debnath's (1962) *Hada Bhoda Samagra* (p. 31)

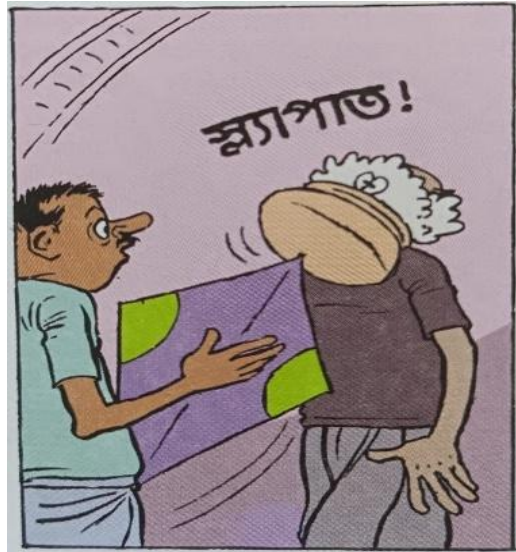


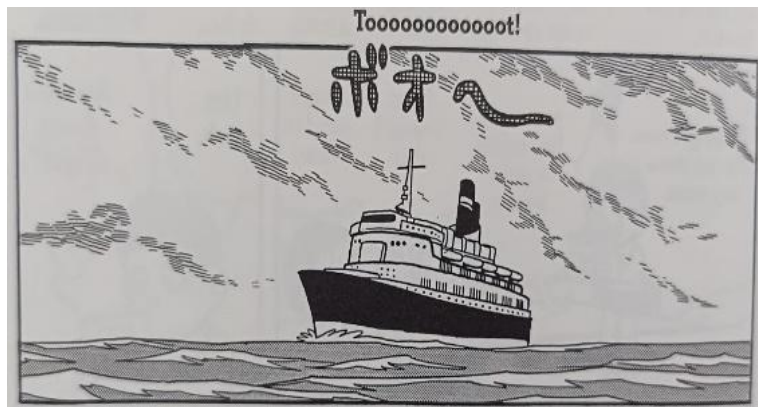
Figure 74. Fujio's (1974) *Doraemon- Volume 4* (p. 130)



Figure 78., represents an eerie auditory picture as *Sunio* holds a snake in his hand, though it appears quite funny with the manga character. The Japanese onomatopoeia of the snake here is /suiru suiru/. The written ideophone on the panel also has scaly features like a snake, adding extra visual effect to it. The sound captures the physical movement of the snake along with its inherent characteristic of stealth and is used to describe the English 'slithering' sound. The reduplication of the word /suiru/ depicts the ongoing action of the snake's movement. In English, we do not find reduplication of words, yet it is translated as 'slither slither' to replicate the Japanese onomatopoeic action, which also does not bring out the original artistic essence of the action.

Figure 79., represents a low, deep noise made by a cruise ship: /bo:/. The /o:/ is quite prolonged and takes a longer duration, representing a deep rumble of engines heard from a distance as it navigates through

Figure 76. Fujio's (1974) *Doraemon- A Selection of Touching Stories* (p. 103)



the sea, bringing out the essence of the visuals. The English translation ‘Toooooooooooooot!’, on the other hand, sounds more like a whistle than a deep hum of the engines. Nevertheless, the prolonged /u:/ sound still brings out the essence of the distant fog horn.

Figure 75. Fujio's (1974) *Doraemon- A Selection of Touching Stories* (p. 100)



Figure 80., depicts a scene where the characters from *Doraemon* are playing baseball when their ball gets stolen by a dog. *Nobita*, on the other hand, clicks a photo with a gadget given by *Doraemon*. The English ‘click’ sound is represented to be /k^hatʃe/, as the Japanese onomatopoeic word for clicking pictures.

The sound also comes with star-like visuals, adding more depth to the ideophone.

Figure 81. describes a scene where *Nobita's* father is shaving. The repetitive action of shaving gives birth to a reduplicative onomatopoeia similar to Figure 78.: /dʒori dʒori/. Even though English does not have reduplication, it is used here as a translation for 'SHBBB SHBBB', which also describes a subtle friction or a rubbing noise. The visual representation of /dʒori dʒori/ gives readers a blade-like impression, evoking the sensation of a blade being scraped across the skin. To get a similar essence, the English translation is written with capital letters.

Figure 77. Fujio's (1974) *Doraemon- Volume 1* (p. 114)



Figure 78. Fujio's (1974) *Doraemon- A Selection of Touching Stories* (p. 046)



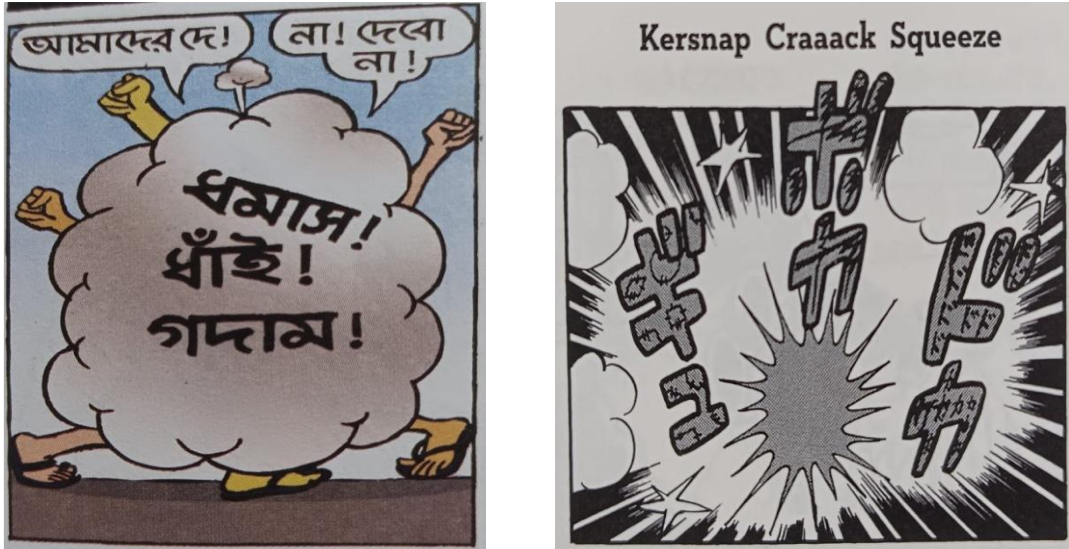
In Figure 82., it is seen that *Nobita* is getting punched right under his chin. This sudden impact also echoes in the air with a forceful Japanese onomatopoeia, /ga:tsə/. This also comes with star-like visuals as it is a common way of showing the speed and force of a punch in visual narratives. Even though the typical

translation for /ga:tsə/ would be 'Pow!', here it is used as 'Kerrrraaaaack' to portray the punch as a lightning strike.

Ideophones or onomatopoeic words are essential in the vibrant world of visual narratives, where pictures and sound effects combine to tell stories. Interestingly, these sounds can differ greatly among cultures, bringing distinct features to comparable scenarios. Similarly,

this study has found certain images in *Hada Bhoda* and *Doraemon* that look similar but have different sounds, making them culture-specific:

Figure 79. Debnath's (1962) *Hada Bhoda Samagra* (p. 15)-Left; Fujio's (1974) *Doraemon- A Selection of Touching Stories* (p. 046)-Right



The above picture, Figure 83., represents two fighting scenes from the Bengali comics, *Hada Bhoda* (left) and the Japanese manga, *Doraemon* (right), respectively. The scene in *Hada Bhoda* looks quite vibrant with colours and expressive artwork. The characters who are involved in the fighting are not visible; only their hands are visible. These characters are within the large bubble-like cloud of dust. This unique visual technique enhances the intensity of the fight while allowing the readers to focus on the onomatopoeic words written on the cloud. There are three very strong ideophones used here depicting the fighting scene: /ḍʰamɔʃ/ /ḍʰāi/ /gɔḍam/. In contrast, the fighting scene in *Doraemon* unfolds with bold, dynamic visuals that look like dusty clouds and lightning strikes to describe its intensity. Similar to the fighting scene of *Hada Bhoda*, characters are not visible; only the three onomatopoeic words bring out the true essence of the fight: /gi:u/ /bɔ:ka/ /ʈ:ka/, giving the essence of a tight grip, a sound of heavy collision, and lightning strike, respectively. So, instead of putting ‘Pow!’, ‘Bam!’ and ‘Crash!’, which are typically used in English comics, the English version goes for a literal translation with ‘Kersnap’, ‘Craaack’ and ‘Squeeze’.

Figure 80. Fujio's (1974) *Doraemon*- Volume 5 (p. 10)- Left; Debnath's (1962) *Hada Bhoda Samagra* (p. 487)-Right

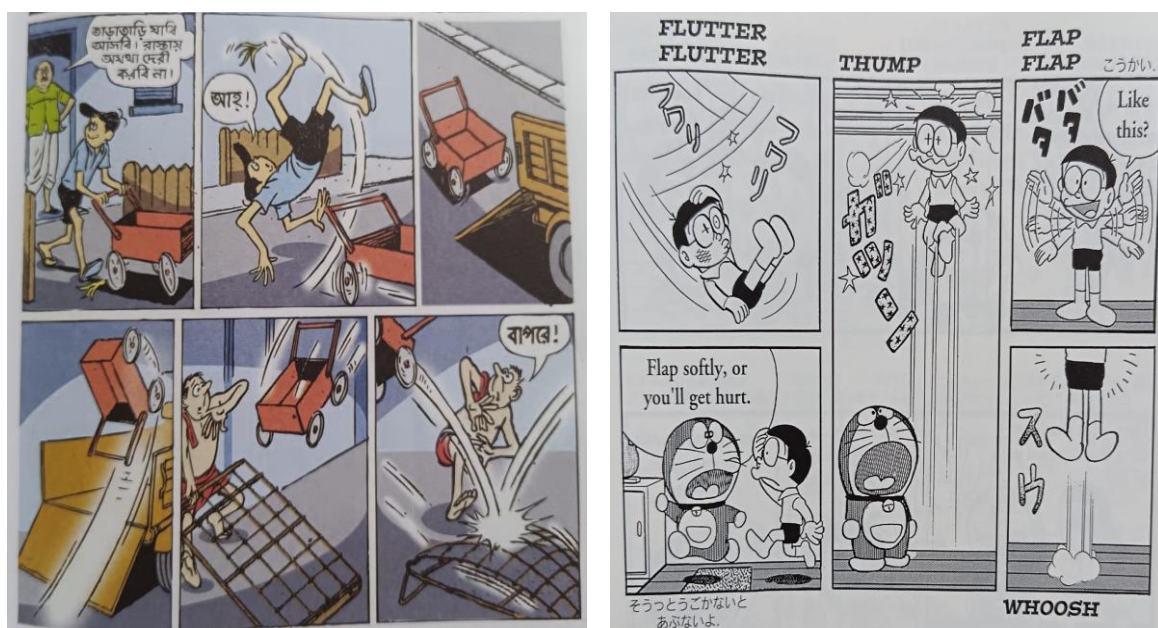


Figure 84. reflects two similar scenes from the Japanese manga, *Doraemon* (left) and the Bengali comics, *Hada Bhoda* (right). The scene from *Doraemon* shows a scary dog's bark that makes *Nobita* dash away in fear. The Japanese onomatopoeia or the ideophone here paints imagery that captures the essence of a dog's bark: /wʌg wʌg/. A typical dog's bark accurately depicts the nasalised sound of the 'a' vowel in Japanese, as well as the nasalised character of the final 'n', making it sound: /wãn wãn/. Here, the presence of the /g/ sound at the end makes the whole panel harsher and stronger to bring out the ferocity of the fierce dog. The translated English version is 'BOW WOW' with a repetitive syllable, 'ow' that brings out the ferocious essence through the bold capital letters. On the other hand, *Hada Bhoda* has a similar situation where *Bhoda* is chased by a less ferocious dog. Yet, *Bhoda*, throwing what he had in his hand, is quite scared as he tries to run away from the bark of the dog: /g^heu/. A typical Bengali comic depiction of a dog's bark would often involve reduplication—such as /g^heu g^heu/—to intensify the sound and convey greater threat or aggression. However, in this scene, the use of a single, non-reduplicated /g^heu/ makes the dog appear subtle and less dangerous, suggesting that *Bhoda's* fear might be exaggerated or comedic in nature.

6.5 Drawing Styles and Multimodality

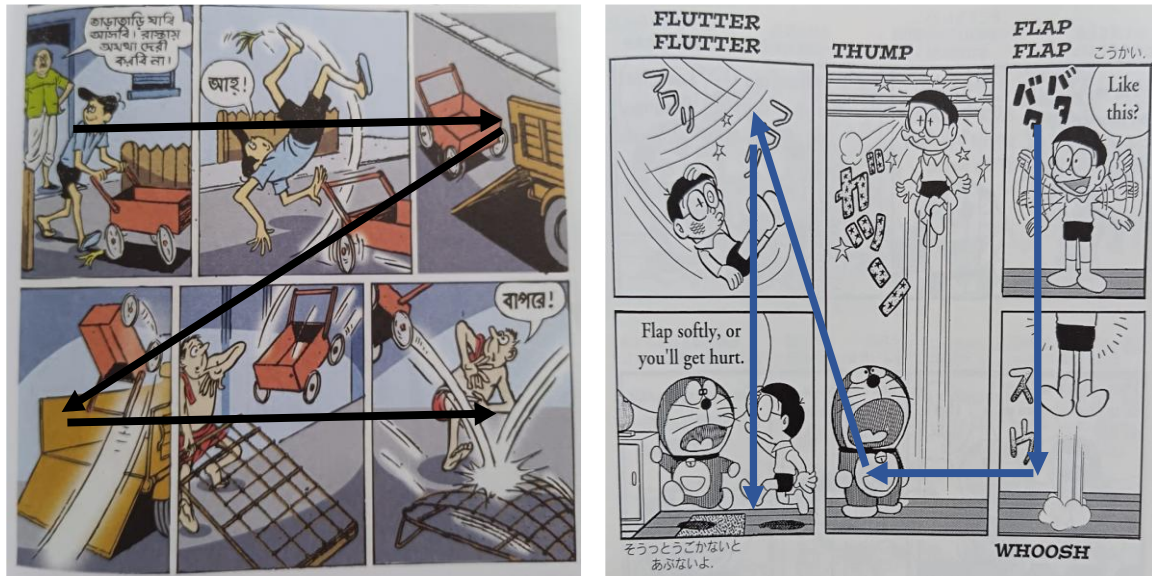
The ideophones also add to the multimodality of visual narratives as the visuals are combined with sound effects to create a more immersive and lively experience for the readers. Even though visual narratives are known for their bimodal forms, the multimodal approach makes the panels come alive with the dynamic interaction between text and art. The readers can navigate through the panels even when there is less or no usage of text:

Figure 81. Debnath's (1962) *Hada Bhoda Samagra* (p. 409)-Left; Fujio's *Doraemon- Volume 4* (p. 046)-Right



The multimodality of these two visual narrative scenes in Figure 85., from *Hada Bhoda* (left) and *Doraemon* (right), relies heavily on the visual elements and cues to convey meanings, emotions, and actions since the presence of textual elements is very minimal. By using a combination of visual cues like symbols, character movements, and imagery, the readers can navigate through the panels to comprehend the embedded meanings. The arrows in Figure 86. indicate the reading paths for the two visual narratives, guiding the readers sequentially through the panels. In *Hada Bhoda*, as mentioned earlier, the navigation follows the usual Z-path; whereas, in *Doraemon*, the navigation is more complex:

Figure 82. Debnath's (1962) *Hada Bhoda Samagra* (p. 409)- Left; Fujio's (1974) *Doraemon- Volume 4* (p. 046)-Right



Navigating through Bengali comics and Japanese manga involves different techniques, as mentioned earlier through the findings of the first objective (see Figures 6. and 7.). Even though the grammar of navigation or the techniques are different, both have a similar commitment to visual storytelling that crosses cultural borders. The distinction in navigation reflects broader cultural approaches to visual storytelling. The Bengali comics, *Hada Bhoda*, prioritises accessibility through the common Western storytelling structure, while the Japanese manga, *Doraemon*, illustrates a more film-like or cinematic and emotive style.

Both *Hada Bhoda* and *Doraemon* provide a rich, immersive experience that merges art and text in unique and complementary ways with distinctive, prominent, and different drawing styles and techniques. While these two visual narratives differ in many ways, they also converge in depicting a few facial expressions. The similarity suggests an intuitive, shared visual language for emotional cues:

Figure 83. Similar Expressions



Source: Debnath's (1962) *Hada Bhoda Samagra* and Fujio's (1974) *Doraemon*

The expression of spewing water out of one's mouth and shouting 'yuck' in response to terrible food is almost done similarly around the globe by different people. *Hada Bhoda* and *Doraemon* from the two cultures of West Bengal and Japan, respectively, brought out this intriguing example (in Figure 86.) of how visual and linguistic aspects interact together to convey meaning. Considering Austin's (1959) speech act theory, these behaviours are not only illustrations but rather communicative or performative acts that convey these characters' disapproval and distaste. Spitting out water and saying, 'yuck' are explicit verbal behaviours that indicate an evaluation of the food's quality.

From the perspective of multimodality, these visuals, along with the texts, increase the whole narrative impact. The visual element, which depicts a persona with exaggerated facial emotions and the act of spitting, delivers an instant representation of disgust. The texts

present, emphasise the reaction while also adding an audio dimension in the readers' or viewers' mind, making the experience more vivid and real.

Throughout this dissertation, multimodality has played a vital role in the images from the two visual narratives of *Hada Bhoda* and *Doraemon*. The colours (in case of *Hada Bhoda*), textures, and line styles of the images evoke sensory associations, making the experience of 'viewing' feel almost tangible. The sound effects, motion lines, and even the exaggerated facial expressions invite the readers to hear, feel, and emotionally connect to the drawn scenes, respectively. Therefore, multimodality with its multimodal and multisensory elements makes the images very interactive, enriching the overall experience of the readers.

The use of modes—visual movements, facial expressions, and vocal utterances- showcases the power of multimodality in visual narratives. It enables readers/viewers from different cultures to comprehend the characters' situations, despite potential cultural differences. The universality here demonstrates how multimodal communication breaks down linguistic barriers and uses human experiences to effectively transmit emotions and actions.

6.6 Discussion

The examples presented in this chapter, drawn from the visual narratives of *Hada Bhoda* and *Doraemon*, provide evidence in support of Austin's (1959) speech act theory. Locutionary acts in visual narratives can be seen as literal portrayals of the speech or the text that appears in the panels or the story that we create in our minds through visuals. It involves the linguistic components of the speaking act, such as the characters' choice of words, sentences and phrases. In visual narratives, locutionary behaviour is often found in speech bubbles, captions or text boxes. The uses of words along with facial expressions and gestures conveying intents, attitudes, demands, promises, suggestions, and other communicative functions can be seen in visual narratives as illocutionary acts; facial expressions, body language and other visual components help to convey meaning. Perlocutionary acts in visual narratives can be

seen through the responding characters in conversation or readers' reactions on some emotional, cognitive or behavioural level. The readers' interpretations and comprehension of the speech act—and how it affects their perception, involvement, or reaction to the narrative—support the notion of the perlocutionary act in visual narratives (Chattopadhyay, 2024).

Through the findings shown in this chapter, the research attempts to show that Austin's (1959) speech act theory plays a crucial role when it comes to visual narratives; it allows a more holistic and comprehensive understanding of the language and meaning-making processes of visual narratives in the light of Linguistics. The viewers' subjective experiences with the narratives are influenced by the visual perspectives used in the picture stories. The findings delve into how Austin's speech act theory enriches our comprehension of visual narratives through emotions, ideophones, and different drawing techniques, revealing their universal characteristics, transcending cultural and linguistic boundaries, and engaging readers with techniques such as visual metaphors, symbols, and multisensory communication.

The art of visual storytelling includes facial expressions, ideophones, and various drawing styles. They collaborate to produce rich, engaging narratives that readers can connect with. Comprehension and appreciation of these components allow the readers to obtain a better understanding of cultural nuances and creative intricacies that contribute to the different emotions, expressions, and forms as depicted in the Bengali comics, *Hada Bhoda*, and the Japanese manga, *Doraemon*. On the other hand, multimodality in visual narratives, enhanced by cultural specificities, provides an even richer and varied storytelling experience. Readers can obtain a better grasp of the cultural backgrounds from which these stories emerge by paying attention to the nuances in language, graphic features such as emotions, expressions, and gestures, onomatopoeias, panel layouts, and other thematic contents. Whether through the vivid art expressions of the *Hada Bhoda Samagra* or the exaggerated delight of the

Doraemon series, these styles, forms, and techniques bring the narratives to life in vibrant and striking ways.

The application of Austin's (1959) speech act theory in the analysis of visual narratives—*Hada Bhoda* and *Doraemon* allows for a more nuanced exploration of how meaning is constructed, negotiated, and experienced by readers. Beyond the traditional linguistic framework, the speech act theory in visual narratives invites readers and viewers to look at images not as static representations but as dynamic communicative acts embedded in a broader cultural and emotional context. Each image, gesture, or visual cue can be interpreted as an act with intended effects, whether it is to inform, persuade, comfort, or entertain.

Visual narratives are not merely passive texts that readers decode; they are interactive semiotic fields in which meaning is actively co-created. For example, in *Hada Bhoda*, the slapstick humour and exaggerated expressions are not just illustrations of action but also function as performative acts that elicit laughter, discomfort, or recognition. Similarly, in *Doraemon*, *Nobita*'s often tearful pleas and *Doraemon*'s futuristic gadgets serve not just narrative purposes but also act as illocutionary cues that deliver moral lessons or provoke emotional empathy. These visuals function like a language, where their meanings rely on the context, the viewer's background knowledge, and an understanding of cultural nuances.

Moreover, these visual speech acts are not isolated within the narrative; they extend outward, interacting with readers' emotions and worldviews. This can be seen as a form of perlocutionary force, where a reader might feel empowered, reassured, or even challenged by what is represented. In this way, visual narratives serve a dual role, depicting a fictional world while simultaneously engaging with real-life social experiences and psychological responses.

What makes this especially compelling is the way they make use of silence, gesture, and implied meaning. Not every panel needs text; often, the absence of words is what delivers the

most potent message. A silent frame of a character looking away, a pause between action sequences, or a subtle shift in colour or composition can serve as a powerful speech act in itself. These acts function on a symbolic level and are often deeply embedded in cultural codes—codes that the readers must learn to interpret through repeated exposure and cultural literacy.

This leads us to consider the role of embodiment in speech acts within visual narratives. Traditional speech act theory centres on spoken or written language, whereas visual storytelling integrates bodily movements, spatial arrangements, and visual rhythms as core communicative tools. Facial expressions, for instance, can reveal sarcasm, sincerity, surprise, or fear without a single word. Likewise, posture and movement within the frame create an embodied grammar and functionality that the readers intuitively learn to interpret. This embodied interaction enhances the depth of illocutionary force in visual narratives.

Furthermore, the reader's cultural background and familiarity with certain visual tropes significantly influence how speech acts are received and interpreted. A gesture that may be humorous or benign in one culture could be read as offensive or serious in another. Therefore, the perlocutionary effects of visual narratives are not universal; they are culturally conditioned and contextually dependent. In *Hada Bhoda*, the use of local dialect and cultural references reinforces a shared context among Bengali readers, whereas in *Doraemon*, the use of futuristic gadgets and stylised emotive visuals align more with Japanese sensibilities and contemporary technological imagination. This cultural specificity in how speech acts operate visually invites scholars to explore cross-cultural semiotics and reader-response theory as complementary methodologies.

Furthermore, there's a pedagogical potential in using visual narratives informed by speech act theory. Teachers can help learners engage with narratives beyond surface-level comprehension, encouraging them to identify locutionary elements (what is being said or

shown), illocutionary forces (what the speaker/author intends to do), and perlocutionary effects (how it impacts the audience). In language learning contexts, particularly with younger learners or multilingual classrooms, comics and manga provide accessible, engaging tools for developing interpretative skills and intercultural awareness.

Therefore, the evolution of visual narratives into digital platforms such as webtoons, memes, and interactive comics has added another layer of complexity to how speech acts function. With clickable elements, scrolling narratives, and embedded sound or animation, the scope for illocutionary and perlocutionary acts is expanding. These multimodal texts enable readers to experience narratives more immersively, responding not just to visuals and text, but to motion, timing, and interactivity. By substituting a dynamic interface for a static page, this layered experience pushes the limits of traditional storytelling. Multimodality, in this sense, doesn't just enrich the narrative—it transforms it, offering readers multiple sensory entry points. As technology evolves, so too does our understanding of what it means to ‘read’ a story. This development opens avenues for further research into how speech act theory can be adapted to digital storytelling formats and how readers, as both interpreters and participants, actively reshape the communicative landscape.

Understanding visual narratives through the lens of speech act theory reveals a rich interplay between image, text, emotion, and culture. It encourages us to view visual narratives not just as entertainment, but as performative texts capable of evoking, persuading, instructing, and transforming. It underscores the communicative power of visuals in storytelling and supports a broader linguistic inquiry into how we make meaning in multimodal, multicultural, and mediated ways.

Chapter 7

Final Insights: Cross-Cultural Visual Narratives and their Cross-Linguistic Variations

Where the Journey Ends, Another Begins...

7.1 Summary

Language, according to Chomsky (1957), is something we start to acquire from birth, eventually learning its structures and nuances. Similarly, we also begin to understand visual narratives as we learn to interpret visual cues and decode the narratives. However, this presents an interesting question: do we grasp the grammar and structure of visual stories in the same way we acquire our mother tongues? Visual narratives are not language, but with all the multisensory properties it has, it aligns well with the characteristics of Linguistics. This study explores and bridges the gap among visual narratives, linguistic cognition, and culture, emphasising the role of cross-linguistic visualisation in communication. Visuality engages the brain in such a way that it enhances human comprehension of language and cultural expressions. Drawing on linguistic theories, it is evident from the analyses shown in Chapters 4, 5, and 6 that cross-cultural notions influence the interpretation of visual narratives by focusing on the selected visual narratives of West Bengal (India) and Japan, respectively. This study reveals the intrinsic cognitive capacities and how aesthetic choices and cultural specifications shape human understanding and production of visual narratives, bringing the innate cognitive abilities to the forefront that underpin the comprehension of Visual Linguistics.

7.1.1 The Interplay of Culture in Visual Narratives

The exploration through the journey of this research into cross-cultural narratives of the Bengali comics, *Hada Bhoda* and the Japanese manga, *Doraemon* along with *Patua* and *Emaki* scroll paintings has led me through a rich fabric of narrative or storytelling techniques, their evolution over time, cultural expressions with the help of a few linguistic theories,

revealing the intricate ways of visual narratives that serve as both mirrors and windows into the societies from which they originate. They provide both a reflection to portray the societies, as well as represent views into other societies, offering cross-linguistic notions. This research highlighted the profound interplay between structure, grammar, and functions, along with synchronic and diachronic variations and speech act theory within the selected visual narratives.

The structural and grammatical foundations of visual narratives, which include elements like panels, gutters, and sequencing principles, provide the bedrock of visual storytelling. Nevertheless, it is the functional aspects that make these structural and grammatical features work together to convey ideas, elicit emotions, impart culture-specific values, and help with the overall story progression, making it more film-like in comprehension and, as a result, giving life to the tales. Visual narratives portray a dynamic interaction between form and function through a variety of culturally specific utterances and cues, such as facial expressions, character gestures, and other iconic symbols, making them an effective medium of cross-cultural communication.

Beyond the structural and functional dimensions, these visual narratives also engage with universal cognitive functions that transcend cultural subtleties to serve as a fundamental way of human communication, comprehension, and social interaction. This study also explored the selected visual narratives diachronically and synchronically and has drawn parallels between the evolution of picture stories and the development of linguistic values like languages in visual narratives. Just as languages have created complex and sophisticated frameworks across different cultures, the properties of visual narratives, too, have enhanced their sequential approaches through panel formations and story progressions over time. The evolution from scroll paintings to modern visual narratives such as comics and manga demonstrates continuity and change, emphasising how narrative structures adapt to cultural and temporal contexts.

Furthermore, the application of speech act theory to visual narratives enriches human comprehension of how these narratives function. Locutionary acts are conveyed through the literal meanings of texts and visuals, illocutionary acts are exhibited through the combination of words, facial expressions, and gestures, and perlocutionary acts are observed as the responses and reactions of both characters present in the visual narratives and their readers. This multimodal perspective of communicative acts within visual narratives highlights the holistic nature of meaning-making processes in visual storytelling.

The findings of this study demonstrated the incorporation of visual narrative grammar, structure, and functions along with its temporal variations and the integration of communicative acts through facial expressions, ideophones, and drawing techniques that produce rich, engaging stories resonating with readers or viewers on multiple levels. The multimodality of visual narratives, emphasised by culture-specific nuances, provides a deeper and more varied storytelling experience. By paying attention to the nuances of textual and visual features that include language, emotions, expressions, and thematic content, readers or viewers can have a better comprehension of the cultural backgrounds and creative intricacies that shape the visual narratives of *Hada Bhoda* and *Doraemon*, respectively.

The analysis of *Hada Bhoda* and *Doraemon* through the perspectives and viewpoints of linguistic theories and cultural analysis reveals the profound ways in which these narratives bridge cultural divisions and, at the same time, promote universal connections. Visual narratives undoubtedly will remain a vital medium for conveying complex cultural stories and human experiences, as they continue to evolve. The journey of understanding visual narratives has both an end and a beginning that opens new pathways for future research and a deeper appreciation of the very art of visual storytelling.

7.1.2 Final Reflections on Theoretical Framework

In conclusion, this study has highlighted the complex and multifaceted character and depth of Visual Linguistics in visual narratives through an interdisciplinary theoretical framework. The three objectives of this study, which are described in Section 1.3 of Chapter 1, were thoroughly analysed by integrating Chomsky's (1957) TGL, Halliday's (1961) SFL, Saussure's (1916) diachronic and synchronic analyses, and Austin's (1959) speech act theory. This complete and comprehensive approach provides valuable insights into how textual and visual aspects work together to produce captivating stories, evoke variations across temporal dimensions, and study culture-specific emotions and gestures.

The first objective under the general objective examined the structural, grammatical, and functional aspects of visual narratives with the help of the two stimuli, *Hada Bhoda* and *Doraemon*, by applying Chomsky's (1957) TGL and Halliday's (1961) SFL. This enabled a more detailed examination of how visual narratives use structure and grammar to function cohesively and convey meaning across different cultures.

The second objective explored the evolution of visual narratives through the selected stimuli, *Hada Bhoda*, *Doraemon*, and especially through *Patua*, and *Emaki* scroll paintings, respectively, by applying Saussure's (1916) diachronic and synchronic approaches. This analysis enabled the exploration of the temporal transformation of visual narratives, from traditional scroll paintings to modern formats, across cross-cultural settings.

Finally, the third objective, explored the communicative impact of visual narratives through the two stimuli, *Hada Bhoda* and *Doraemon*, by using Austin's (1959) speech act theory. This approach facilitated an in-depth study of how textual and visual components interact within visual narratives, influencing the interpretations of readers and viewers through facial expressions, ideophones, and drawing styles across cultures.

This interdisciplinary approach not only improves and enriches the study of visual narratives but also sheds light on broader issues in visual communication and linguistic philosophy. The study advances and contributes to a deeper comprehension of the grammar and functions of visual narratives, their temporal variations, and their ability to communicate complex culture-specific emotions and gestures by investigating the nuanced interplay between visual and linguistic components.

7.2 Findings in Congruency with the Objectives

The first specific objective of this research explores how the visual narratives of *Hada Bhoda* and *Doraemon* reflect distinct cultural perspectives through their structural organisation, visual grammar, and functional aspects that are socially embedded within them. The analysis reveals that each visual narrative's panel sequencing and layout reflect unique cultural storytelling conventions, with specific visual orderings and arrangements contributing to their cultural relevance. Particular illustration styles, narrative frameworks, and panel structures demonstrate how the respective Bengali and Japanese cultures interpret and arrange visual storytelling in their own ways. Social contexts illustrated through themes like family values, friendly exchanges, school conversations, sportsmanship, and culinary culture further distinguish the visual narratives and bring out cultural values and norms through functional aspects of Linguistics. The visual language in both of these visual narratives relies on culturally specific expressions, gestures, and symbolic elements, emphasising the idea that structure, grammar, and function work together to express and communicate linguistic identities and values in culturally and visually distinct forms.

The second specific objective of this research analyses how the structures of traditional *Patua* and *Emaki* scroll paintings differ from the visual grammar found in contemporary Bengali comics, *Hada Bhoda*, and Japanese manga, *Doraemon*. The findings convey that the evolution from scroll paintings to respective Bengali comics and Japanese manga panels reflects broader shifts in narrative techniques and cultural storytelling methods. The iconic,

achronological panel arrangements and the natural gutters in scroll paintings have transformed over time, thereby adapting to new narrative structures that highlight panel structuring in culture-specific manners to aid comprehension. This shift from scroll painting formats to strictly bound segmented panels in contemporary visual narratives aligns with changes in how events are selected and illustrated, along with social contexts that have impacts on both the choice of scenes and their portrayals. Such diachronic shifts in visual narratives reflect the development of language and society, where complexity in visual structure and grammar emerges to bring out meaning and engagement. On the other hand, synchronic analysis of these visual narratives reveals both cultural uniqueness and universal storytelling factors through cross-linguistic analysis. This cross-linguistic analysis shows that visual narratives have undergone significant change, yet they retain their core storytelling aspects, bridging traditional and contemporary cultural narratives through evolving visual structures, grammar and functions.

The third specific objective of this research examines how visual culture is communicated in the narratives of *Hada Bhoda* and *Doraemon* through the use of communicative acts, facial expressions, ideophones, and distinct drawing styles. The analysis emphasises the speech act theory, and the findings reveal how visual elements serve as locutionary acts depicting literal representations that align with textual components, while facial expressions and gestures function as illocutionary acts, conveying deeper comprehensions and emotional nuances. The reactions of the characters present in the visual narratives and the responses of the audience according to how they view and engage with the visual narratives represent the perlocutionary acts on emotional and cognitive levels. The analysis emphasises that facial expressions play crucial roles in conveying emotions and character intentions by significantly enriching the visual narratives' communicative powers. Similarly, ideophones evoke sensory experiences in the readers and provide notable differences in their usages, cross-linguistically, across the two cultures. Furthermore, the drawing styles in these visual

narratives illustrate unique artistic traditions contributing to the overall holistic and multisensory experience of visual storytelling. Therefore, this analysis highlights the different communicative aspects and enriches the storytelling techniques in both *Hada Bhoda* and *Doraemon*, enabling the readers to engage in appreciating cross-linguistic nuances. The analysis, therefore, unveils the multimodal nature of visual narratives by demonstrating how they dive into the notions of understanding cultural nuances while also showcasing universal storytelling techniques that transcend linguistic boundaries.

7.3 Concluding Reflection and Implications of the Study

In light of the exploration presented in this dissertation, it is clear that the true essence of things and events is not intrinsic but shaped by cultural perceptions. This dissertation delved into the implications of cross-cultural visual narratives through the cross-linguistic analysis of the Bengali and Japanese cultures with the help of the selected stimuli, which reveal how they represent different cultures and create a unique and distinct communication system within a universally accepted medium.

The examples taken in this dissertation illustrate the concept of cross-linguistic variations in visual narratives by emphasising the importance of cultural context in representation. The research problem proposed that cross-cultural differences stem from the visualisation of things and events not based on their intrinsic essence, but on how they are perceived and represented across cultures. This viewpoint rejects the notion of a single, underlying meaning and instead emphasises the dynamic interplay between visual and linguistic components in conveying diverse culture-specific meanings and comprehension.

Furthermore, the interdisciplinary approach enriches the study of visual narratives and sheds light on broader challenges in the study of visual communication and linguistic philosophy. It highlights the importance of cultural context in shaping visual narratives and eventually offers insights for future research in Visual Linguistics and related fields. The study of cross-

cultural variations revealed the profound impact of cultural influences on the creation, communication, and interpretation of picture stories, highlighting the requirement and necessity of considering these variations in the study of visual communication.

In conclusion, this dissertation has come to the point that visual narratives are not static or universal representations but dynamic cultural constructions, deeply embedded in specific linguistic and communicative frameworks. By employing the conceptual framework already discussed in Chapter 1, grounded in foundational linguistic theories, this research translated abstract linguistic insights into a practical analytical model. Through the systematic analysis of *Hada Bhoda*, *Doraemon*, and the *Patua* and *Emaki* scrolls, the study has illuminated how visual and linguistic elements operate as culturally embedded tools of communication. As a pragmatic researcher, I recognise that the value of such a framework lies not only in its theoretical depth but also in its adaptability to real-world analyses of visual media. The intersection of visuals, textual elements, and cross-linguistic comparison revealed how ideologies, temporalities, and communicative norms are constructed and interpreted differently across cultures. This research thus advocates for a shift toward culturally responsive methodologies in Visual Linguistics and communication studies that are attuned to variation, context, and evolving modes of meaning-making. By bridging disciplinary boundaries, the study contributes to a growing body of research that views visual storytelling as both a linguistic practice and a site of cultural negotiation, offering new directions for scholars, educators, and media practitioners alike.

Through systematic analysis, this study has shown how ideologies, temporal structures, and communicative norms are differently constructed and interpreted across cultures, even when visual grammar and narration are assumed to be ‘universal’. The interdisciplinary approach adopted here opens up broader perspectives for the fields of Visual Linguistics, language education, and visual communication studies. As such, several key implications emerge from this research—both theoretical and applied in nature.

7.3.1 Theoretical Implications

This dissertation contributes to the evolving field of Visual Linguistics by looking at visual narratives through a cultural and cross-linguistic viewpoint rather than treating them as universal forms. Therefore, it expands its analytical scope as a result of this. The findings challenge the assumption of universal meaning in visual storytelling, reinforcing the theoretical position that language and image are not neutral transmitters but culturally encoded systems of representation. The research supports a multimodal approach grounded in the fields of pragmatics, semiotics, and sociolinguistics, highlighting how visual narratives reflect ideological positioning, communicative norms, and culturally specific conceptual metaphors. Therefore, this research suggests that we need pluralistic theoretical models that are sensitive to local forms of expression and meaning-making in visual texts. This very conclusion is repeatedly confirmed through the findings of this dissertation that visual meaning does not travel easily across cultural contexts.

7.3.2 Methodological Implications

This dissertation underscores the value of a qualitative, comparative, and context-sensitive methodology in the study of visual narratives. While the study primarily employed content analysis, it also incorporated field-based inquiry, particularly in the second objective, where a short visit to *Naya, Pingla* was undertaken to directly observe *Patua* scroll paintings and to conduct an interview with *Ranjit Chitrakar*, one of the key artists. This blend of textual and field data strengthened the empirical grounding of the research, enriching the analysis with first-hand cultural insight. The overall methodological design integrates visual grammar, narrative sequencing, cross-linguistic comparison, temporal factors, the study of ideophones, facial expressions, other visual cues, and cultural symbolism—elements often underexplored in multimodal discourse research. This triangulated approach facilitated a nuanced understanding of the artefacts while preserving cultural specificity. Furthermore, looking at these aspects collectively made it possible to develop a more layered comprehension of the

materials while keeping their cultural context intact. Therefore, combining visual, oral, and written traditions suggests a way forward for future researchers who may wish to build such flexible yet thorough methods that could suit a range of cultural and narrative settings.

7.3.3 Pedagogical Implications

The findings can also offer meaningful insights for educators, particularly in the fields of literacy, language education, cultural studies, and media pedagogy. By treating visual narratives as culturally grounded texts, educators can encourage students to engage in critical visual literacy, appreciating how culture shapes storytelling, symbolism, and interpretation. The inclusion of *Hada Bhoda*, *Doraemon*, *Patua*, and *Emaki* picture stories provides a template for incorporating non-Western and traditional art forms into academic discourse, promoting more inclusive and decolonised curricula. Furthermore, the study suggests pedagogical value in comparing visual narratives cross-culturally, enabling learners to recognise the diversity of representational systems and enhancing their intercultural communicative competence. It advocates for integrated multimodal instruction that draws from visual, oral, and textual modes to build holistic understanding among learners. Because people naturally draw on multiple modalities simultaneously when interpreting picture stories, working with visual narratives provides valuable insight into how language functions in real-time. It therefore illustrates that meaning is not created by words alone but through the interplay of different semiotic resources as they unfold. In this sense, Visual Linguistics offers a strong framework for exploring the sequential, layered, and multimodal nature of human language processing. According to Chattopadhyay (2025), the structural and functional features found within the visual narratives can support students' metacognitive skills, which are crucial in pedagogy and also are relevant in language classrooms. Using such visual narratives in teaching can allow educators to address both the scientific aspects of language and the artistic qualities of literature in a balanced and meaningful way.

7.4 Expanding Horizons and Final Thoughts

In this dissertation, the Bengali comics, *Hada Bhoda* and the Japanese manga, *Doraemon*, along with Bengali *Patua* and Japanese *Emaki* scroll paintings were selected, respectively, to explore the intricate interplay between visual and linguistic components within the notion of cross-cultural visual narratives. These cultural nuances have provided a rich panorama for analysing the dynamic and multifaceted characteristics of Visual Linguistics. Nevertheless, the scope of this research expands far beyond the Bengali and Japanese cultures, only. In future studies, many other cultures and societies can be explored and analysed under the same or different theoretical framework of Linguistics, which could reveal even more about how visual narratives operate across diverse cultural contexts. This dissertation explores the visual and linguistic dimensions of cross-cultural storytelling through the comparative study of Bengali comics *Hada Bhoda* and Japanese manga *Doraemon*, alongside Bengali *Patua* scroll paintings and Japanese *Emaki* scrolls. These specific examples were chosen not only for their popularity and cultural significance but also because they encapsulate complex layers of tradition, language, humour, temporality, and moral education embedded within their respective societies. By analysing these four visual narrative forms, the dissertation delves into how different societies use visual language to articulate cultural values, norms, and social realities. Through the framework of Visual Linguistics, this work dissects the grammar of images and their interaction with text, emotion, sound (through ideophones/onomatopoeia), and motion, to shed light on how meaning is visually constructed and interpreted across cultures.

The study of visual narratives, especially those rooted in traditional and popular forms like comics and scroll paintings, offers unique insights into the sociocultural underpinnings of storytelling. While my research focuses on the visual grammar, ideophones, drawing styles, and thematic content of *Hada Bhoda* and *Doraemon*, it becomes evident that such narratives are not merely for entertainment—they are cultural texts, deeply embedded with values,

ideologies, and societal reflections. When we analyse them through a cross-linguistic and cross-cultural lens, we uncover how societies imagine childhood, education, discipline, humor, and even resistance differently. These differences, however, do not exist in isolation. They open up possibilities for larger dialogues between communities and cultures, and allow us to better understand how visual semiotics operates in diverse sociopolitical settings.

Art has political implications, but socially conscious art actively protests against war, injustice, and corruption. This kind of art is a relatively recent phenomenon, emerging with power and urgency in the early nineteenth century through the works of Francisco Goya (Caplow, 2009). While many artists have painted images with political intent, printmaking has provided a particularly powerful medium of expression. Because of its reproducibility, affordability, and graphic potency, the print lends itself well to protest and public engagement. It moves beyond the confines of galleries into public spaces, circulating ideas and stirring collective sentiments. Political prints have always been intended to engage the viewer, change opinion, and inspire action.

In the twentieth century, waves of printmaking around the world—across Europe, the U.S., China, and Mexico—reflected the urgency to communicate with mass audiences. Printmakers often worked collectively to produce vivid, raw, and timely representations of social realities (Caplow, 2009). However, the medium of printmaking has traditionally been sidelined in mainstream art discourse. Caplow further notes that artists like Rembrandt and Goya, though renowned printmakers, are celebrated primarily for their paintings. Others, such as Käthe Kollwitz, who centered their work in print, remain under-recognised. Because prints are often ephemeral—pasted on walls, handed out as leaflets, or banned for their content—they exist on the margins of canonised art history. Political art has often been dismissed as propaganda, especially in U.S. discourse post-1940s, diminishing its critical value and excluding it from major exhibitions and academic texts.

The link between this legacy and my dissertation work lies in the ways visual narratives continue to function as tools of cultural communication, subtly or overtly conveying ideological messages. *Hada Bhoda* and *Doraemon* are not just children's comics; they mirror the societal, educational, and moral expectations of their respective cultures. As a result, they encode the political climates in which they were created and circulated. Much like political prints, these narratives are meant to be seen and understood by broad audiences. They shape values, behaviours, and perceptions through visual-linguistic strategies. The study of such texts within Visual Linguistics opens up new horizons—inviting us to interrogate power, identity, resistance, and cultural continuity in visual forms. As socially engaged art pushes boundaries through print, film, and posters, comics and scrolls likewise carry the seeds of cultural commentary and potential resistance.

By analysing the linguistic variations in the selected visual narratives, a better understanding of the basic principles of picture stories has been obtained, along with the distinct ways in which different cultures interpret and convey meanings and comprehension. This ongoing exploration can promise to further enrich our comprehension of Visual Linguistics and its profound influence on global communication. There exists a vast, largely untapped field of visual storytelling traditions from numerous other regions that could provide equally rich material for study. For example, Indigenous Australian dot paintings, Native American ledger art, African comic traditions, Latin American mural narratives, or Buddhist and Nepali Thangka paintings each contain deeply symbolic visual languages shaped by their socio-historical and environmental contexts. These traditions often incorporate unique perspectives on time, space, community, and spirituality, revealing alternate ways of constructing and perceiving reality through visuals.

From a pragmatic research perspective, the application of Visual Linguistics to a wide range of visual traditions offers fertile ground for interdisciplinary inquiry and methodological innovation. While this dissertation focuses on the cross-linguistic variations of Bengali

comics such as *Hada Bhoda* and Japanese manga like *Doraemon*, along with the respective cultural scroll paintings, the potential for broader application of similar frameworks is evident. Visual storytelling—across comics, scroll paintings, digital media, or community murals is never neutral; it is deeply embedded in specific socio-cultural, historical, and even political contexts. As such, using Visual Linguistics to dissect these forms can yield insights not only into semiotic structures but also into how communities express, remember, and resist.

One promising avenue for future research lies in post-conflict visual narratives, where picture stories operate not merely as vessels of personal testimony but as politically charged texts that reflect and shape the collective memory of a society recovering from war. In the aftermath of conflict, when wounds are raw and words may fail, the visual form becomes a potent medium for articulating trauma, grief, and fragmented identity. These narratives are not just about recounting what happened—they deal with how people remember, reinterpret, and live with the consequences of violence and displacement.

Similarly, feminist graphic novels could also be studied to understand how gender roles, domestic life, and bodily autonomy are negotiated in cultures where direct critique might be censored or discouraged. The layering of satire, symbolism, and intertextuality in these works, when viewed through the lens of Visual Linguistics, can reveal subtle yet powerful strategies of dissent and self-expression.

Digital media presents yet another compelling area; the evolution of visual grammar in digital spaces like meme culture, interactive webtoons, or social media comics challenges traditional boundaries. Memes, for instance, rely heavily on shared cultural knowledge, rapid circulation, and punchy visuals, often bending or breaking linguistic rules to heighten impact. In this context, Visual Linguistics must not only adapt in methodology but also in its scope, to account for the co-creation of meaning between the content creator and audience. Here, the

lines blur between author and viewer, fiction and lived experience, fixed narrative and open-ended engagement.

In short, Visual Linguistics, when applied flexibly and contextually, can provide powerful insights into how communities speak visually, about their histories, politics, identities, and aspirations. While my dissertation engages deeply with Bengali and Japanese traditions, I see it as a stepping stone toward a much broader, more inclusive exploration. There is immense potential to expand this line of inquiry into indigenous visual traditions, diasporic comics, and even protest art, all of which call for nuanced and context-aware visual linguistic analysis.

The qualitative content analytical approach provided valuable insights into how visual language operates across cultures, cross-linguistically, and intentionally excludes audience responses to maintain focus on the inherent characteristics of the visual narratives. Altogether, this research contributes to the study of Visual Linguistics by exploring how visual storytelling transcends and highlights the significance of visual media in conveying cultural values and norms, thereby setting the stage for future researchers on cross-linguistic variations in visual culture.

As McCloud (1993) rightly points out, ‘The power of comics lies in its ability to blend visual and textual elements to convey meaning’, it allows a multilayered and multimodal experience where visuals provide immediate context and emotion, while text offers more detailed information through captions, dialogues and ideophones. Together, they create a unified narrative capable of communicating complicated ideas and cultural nuances in an interesting and accessible way. For example, in the field of Media, this research can offer valuable insights into how visual storytelling transcends cultural boundaries. As a result, the knowledge provided by this study can inform the creation of content that resonates across cultures, allowing media producers to craft stories that are both globally accessible and

culturally sensitive. Also, in the field of Psychology, this study offers insight into the ways visual storytelling influences both cognitive and emotional responses. Such an understanding can assist psychiatrists, psychologists, and psycholinguists in examining and exploring how cultural background shapes emotional interpretation and cognitive growth. Moreover, most importantly, this research has significant implications for future generations in the realm of Education. In the age of AI, visual storytelling can improve traditional teaching strategies and utilise the power of visual narratives. Incorporating such diverse visual narratives into curricula can also enhance creativity and engagement, ultimately paving the way for a more holistic and culturally responsive approach to education, benefitting both the current and the future generations...

References

- Adhikari, S. (2022). *The 10 Oldest Ancient Civilizations That Have Ever Existed*. Ancient History Lists. <https://www.ancienthistorylists.com/ancient-civilizations/10-oldest-ancient-civilizations-ever-existed/>
- Ahire, D. (2020). *Endless Scroll – The Genesis of Bengal’s Pattachitra Art*. The Sarmaya Guide to Bengal Pattachitra. <https://sarmaya.in/guides/endless-scroll-the-genesis-of-bengals-pattachitra-art/>
- Akter, R., Thilsted, S. H., Hossain, N., Ishihara, H., & Yagi, N. (2019). Fish is the Preferred Animal-Source Food in the Rural Community of Southern Bangladesh. *Sustainability*. <https://doi.org/10.3390/su11205764>
- Alambrouk. N. A. (2020). Understanding Derrida’s “Structure, Sign, and Play”. *English Language Teaching and Linguistics Studies*. [DOI:10.22158/eltls.v2n4p43](https://doi.org/10.22158/eltls.v2n4p43)
- Alexander, A. (2023). Literature and Its Uses. *The Joy of Learning: Engaging Faith and Scholarship Vol. 3. Essays in Humanities, Law and Education*. https://www.academia.edu/111238639/Literature_and_Its_Uses?uc-g-sw=35945646
- Almurashi, W. A. (2016). An introduction to Halliday’s systemic functional linguistics. *Journal for the Study of English*, 4(1), 70-80. <https://doi.org/10.5296/jsel.v4i1.9423>
- Alves, T., Simões, A., Figueiredo, R., Vala, M., Paiva, A., & Aylett, R. (2008). So tell me what happened: Turning agent-based interactive drama into comics. *AAMAS* (3) (pp. 1269-1272). <https://www.researchgate.net/publication/228359599>
- Andrade, C. (2020). Learning curve: The Inconvenient Truth about Convenience and Purposive Samples. *Indian Journal of Psychological Medicine*, 42(2), 169–171. [DOI: 10.1177/0253717620977000](https://doi.org/10.1177/0253717620977000)

- Anwar, S. (2012). *A Pragmatic Analysis on Directive Utterances Used in Donald Duck Comics Manuscript*. Unpublished Research Paper. Surakarta: Muhammadiyah University of Surakarta.
- Ara, R. (2006). Wittgenstein's concept of language games. *Al-Hikmat*, 26, 47–62.
- Austin, J. L. (1959). *How to do Things with Words*. The William James Lectures: Harvard University.
- Baird, C. S. (2020). Do blind people dream in visual images? *The Top 50 Science Questions with Surprising Answers*. <https://www.wtx.edu>
- Bates, E. (1976). *Language and context: The acquisition of pragmatics*. Academic Press.
- Bavali, M., & Sadighi, F. (2008). Chomsky's universal grammar and Halliday's systemic functional linguistics: An appraisal and a compromise. *Journal of Pan-Pacific Association of Applied Linguistics*, 12(1), 11-28. <https://www.studocu.com/es-ar/document/universidad-nacional-de-salta/ingles-i/linguistics-yule-y-chomsky/90242132>
- Belal, H. M., Shirahada, K., & Kosaka, M. (2013). Value Co-creation with Customer through Recursive Approach Based on Japanese Omotenashi Service. *International Journal of Business Administration*. <http://dx.doi.org/10.5430/ijba.v4n1p28>
- Bell, M. M. (1998). The Dialogue of Solidarities, Or Why the Lion Spared Androcles. *Sociological Focus*. <https://doi.org/10.1080/00380237.1998.10571100>
- Bender, L. (1944). The psychology of children's reading and the comics. *The Journal of Educational Sociology*, 18(4), 223–231. <https://doi.org/10.2307/2262695>
- Bloch, B. & Trager, G. L. (1942). *Outline of linguistic analysis*.

- Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.
- Bramlett, F. (2012). Introduction. In: Bramlett, F. (eds) *Linguistics and the Study of Comics*. Palgrave Macmillan, London. https://doi.org/10.1057/9781137004109_1
- Brandom, R. (2011). *Perspectives on pragmatism*. Harvard University Press.
- Bruner, J. (1975). The Ontogenesis of Speech Acts. *Journal of Child Language* 20(1), 1–19.
- Bryman, A. (2004). (2nd ed). *Social Research Methods*. Oxford: Oxford University Press.
- Bury, S. J., & Mellon, A. W. (2022). *British visual satire, 18th- 20th centuries*.
<https://sites.google.com/mtroyal.ca/wordsandpictures/references>
- Byrne, S. (2012). *Ludwig Wittgenstein's legacy to cognitive psychology: Concepts as participatory* (Doctoral dissertation). National University of Ireland, Maynooth.
- Cacioppo, J. T., & Petty, R. E. (1982). The Need for Cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>
- Capps, J. (2019). *Pragmatic theory of truth*. In E. N. Zalta, *The Stanford encyclopedia of philosophy* (Fall 2019 Edition). Stanford University.
<https://plato.stanford.edu/entries/truth-pragmatic/>
- Chattopadhyay, P. (2024). Speech Act Theory in Visual Narratives: An Analysis of Communication through Sequential Art. *Gipan*, 6(1), 119–133.
<https://doi.org/10.3126/gipan.v6i1.68143>
- Chattopadhyay, P., Subedi, A., & Sharma, S. R. (2024). Teaching with visual narratives: A practical usage of English language. *Bodhi: An Interdisciplinary Journal*, 10(3), 112–125. <https://doi.org/10.3126/bodhi.v10i3.76465>

Chōjū-jinbutsu-gig. (2024). *Wikipedia*.

<https://en.wikipedia.org/wiki/Ch%C5%8Dj%C5%AB-jinbutsu-giga>

Chomsky, N. 1977, *Language and Responsibility*, New York: Pantheon Books.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton & Co.

Chute, H. (2016). *Why comics? From underground to everywhere*. Harper.

Cohn, N. (2005). Un-defining 'Comics.' *International Journal of Comic Art*, 7(2), 38-45.

https://www.visuallanguagelab.com/P/NC_Undefining_Comics.pdf

Cohn, N. (2012). Comics, Linguistics, and Visual Language: The Past and Future of a Field (pp. 92-118). *Linguistics and the Study of Comics*. [10.1057/9781137004109_5](https://doi.org/10.1057/9781137004109_5)

Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. Bloomsbury Academic.

Cohn, N. (2014). The architecture of visual narrative comprehension: The interaction of narrative structure and page layout in understanding comics. *Frontiers in Psychology*, 5, Article 680. <https://doi.org/10.3389/fpsyg.2014.00680>

Cohn, N. (2018). Visual language theory and the scientific study of comics. *Empirical comics research: Digital, multimodal, and cognitive methods* (pp. 305-328). London: Routledge. <https://www.visuallanguagelab.com/P/2018.ESiC.NC.pdf>

Cohn, N. (2019). Your brain on comics: A cognitive model of visual narrative comprehension. *Topics in Cognitive Science*, 11(3), 507–526. <https://doi.org/10.1111/tops.12421>

- Cohn, N., & Magliano, J. P. (2019). Editors' introduction and review: Visual narrative research: An emerging field in cognitive science. *Topics in Cognitive Science*, 11(4), 635–652. <https://doi.org/10.1111/tops.12473>
- Cohn, N., Taylor-Weiner, A., & Grossman, S. (2012). Framing attention in Japanese and American comics: Cross-cultural differences in attentional structure. *Frontiers in Psychology*, 3, Article 349. <https://doi.org/10.3389/fpsyg.2012.00349>
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.
- Cummins, J & Davidson, C 2007, *International handbook of English language teaching*, New York: Springer.
- Dan, S. S. (1985) *Marine Fishery of West Bengal Coast. Marine Fisheries Information Service*. <http://eprints.cmfri.org.in/id/eprint/4409>
- Danziger, K. (1988). Review of *Cognition as intuitive statistics* [Review of the book *Cognition as intuitive statistics*, by G. Gigerenzer & D. J. Murray]. *Canadian Psychology / Psychologie canadienne*, 29(4), 370–371. <https://doi.org/10.1037/h0084560>
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray
- De Saussure, Ferdinand. (1966). *Course in General Linguistics (Edited by Charles Bally and Albert Sechehaye, Translated by Wade Baskin)*. New York, Toronto, London: McGraw-Hill Book Company.
- Debnath, N. (1962). *Hada Bhoda*. Sera Handa Bhonda Samagra.

- Dement, W., & Kleitman, N. (1957). The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming. *Journal of Experimental Psychology*, 53(5): 339–46. https://psychologyhacked.org/?page_id=585
- Diep, E. (2019). A brief history of American comics. *MR Comics*.
<https://medium.com/mrcomics/an-introduction-to-comics-in-the-west-part-1-american-comics-f043aaa55781>
- Dummett, M. (1959). Truth. *Proceedings of the Aristotelian Society*, 59(1), 141–162.
<https://doi.org/10.1093/aristotelian/59.1.141>
- Eisner, W. (1985). *Comics & Sequential Art*. Tamarac, FL: Poorhouse Press.
- Ekman, P. (1971). *Universals and cultural differences in facial expressions of emotion*. Paul Ekman Group. <https://www.paulekman.com/wp-content/uploads/2013/07/Universals-And-Cultural-Differences-In-Facial-Expressions-Of.pdf>
- El Refaie, E. (2009). Multiliteracies: How readers interpret political cartoons. *Visual Communication*, 8(2), 181–205. [DOI:10.1177/1470357209102113](https://doi.org/10.1177/1470357209102113)
- Emakimono. (2024). *Wikipedia*. <https://en.wikipedia.org/wiki/Emakimono>
- Ezzat, A. (2021). Animals in Human Situations in Ancient Egyptian Ostraca and Papyri. *MDPI*. <https://doi.org/10.3390/arts10030040>
- Florek, S. (2015). Our Global Neighbours. *Australian Museum*.
<https://australian.museum/blog-archive/science/our-global-neighbours-scroll-paintings/>

- Forceville, C. (2009). Non-verbal and multimodal metaphor in a cognitivist framework: Agendas for research. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal metaphor* (pp. 19–42). Mouton de Gruyter. [DOI:10.1515/9783110197761.5.379](https://doi.org/10.1515/9783110197761.5.379)
- Fraenkel, J. R., & Wallen, N. E. (2006). *How to Design and Evaluate Research in Education*. New York, NY: McGraw-Hill. <https://doi.org/10.4236/crcm.2022.118046>
- Frasco, E., Sherr, L., Dedios Sanguineti, M. C., Vera San Juan, N., & Burgess, R. (2025). The impact of creative arts-based interventions for mental health in conflict-affected contexts: A systematic narrative review. *SSM - Mental Health*, 5, 100419. <https://doi.org/10.1016/j.ssmmh.2025.100419>
- Fujio, F. F. (1969). *Doraemon (Touching Stories)*. Shogakukan English Comics.
- Fujio, F. F. (1969). *Doraemon (Vol 1)*. Shogakukan English Comics.
- Fujio, F. F. (1969). *Doraemon (Vol 2)*. Shogakukan English Comics.
- Fujio, F. F. (1969). *Doraemon (Vol 3)*. Shogakukan English Comics.
- Fujio, F. F. (1969). *Doraemon (Vol 4)*. Shogakukan English Comics.
- Fujio, F. F. (1969). *Doraemon (Vol. 5)*. Shogakukan English Comics.
- Gangopadhyay, U. (2018). Scroll Paintings by the Patua Clan of West Bengal. *Outlook Traveller*. <https://www.outlooktraveller.com/explore/culture/scroll-paintings-by-the-patua-clan-of-west-bengal>
- Genji Monogatari Emaki. (2024) *Wikipedia*. https://en.wikipedia.org/wiki/Genji_Monogatari_Emaki

- GuruNavi. (2017). *Japanese Grilled Fish: The Simple Delicacy that's Surprisingly Easy to Make at Home*. <https://gurunavi.com/en/japanfoodie/2017/05/grilled-fish.html>
- Halliday, M. A. K. (1961). Categories of the theory of grammar. *WORD*, 17(2), 241-292. <https://doi.org/10.1080/00437956.1961.11659756>
- Harari, Y. N. (2015). *Sapiens: A brief history of humankind* (pp. 25-32). Harper: New York, NY.
- He, Y. (2018). *Culinary integration and sweet imagination: The case of Japanese confectionery under globalization* (Master's thesis, University of Oslo). *Department of Culture Studies and Oriental Languages, Faculty of Humanities*. <https://www.duo.uio.no/handle/10852/2855>
- Hobbes, T. (1651) *Leviathan*. London: Printed for A. Crooke. [10.5479/sil.59773.39088001833995](https://www.duo.uio.no/handle/10852/2855)
- IGN. (2012). *Doraemon Creator Dies*. IGN. <https://www.ign.com/articles/1996/10/02/doraemon-creator-dies>
- Ikei, M. (2000). Baseball, "Besuboru, Yakyu": Comparing the American and Japanese games. *Indiana Journal of Global Legal Studies*, 8(1), 73-79. <https://www.jstor.org/stable/20644757>
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.
- Knight, J. (1997). On the Extinction of the Japanese Wolf. *Nanzan University*. <http://www.jstor.org/stable/1178791>

- Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 20(4), 265–276.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC3222570/>
- Kramnick, J. (2018). *Paper Minds: Literature and the Ecology of Consciousness*. Chicago: University of Chicago Press, 2018. 224 pp.
- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage. [ISBN: 9781506395661](https://doi.org/10.1080/1059147042000162892)
- Kumar, A. (2019). The use theory of meaning: A reading of Wittgenstein's *Philosophical investigations*. *Journal of Emerging Technologies and Innovative Research*, 4 pages.
<https://doi.org/10.5281/zenodo.4399324>
- Lemelson-MIT. (n.d.). *John Pemberton*. <https://lemelson.mit.edu/resources/john-pemberton>
- Lenneberg, E.H 1967, *Biological Foundations of Language*, London: Wiley.
- Lent, J. A. (2001). *Illustrating Asia: Comics, humor magazines, and picture books*. University of Hawai'i Press.
- Leon, L. (2017). *The Evolution of Political Cartooning in the New Media Age: Cases from Australia, the USA and the UK*. *Satire and Politics*(pp.163-191).
http://dx.doi.org/10.1007/978-3-319-56774-7_6
- Levin, S. R. (1976). Concerning what kind of speech a poem is. In Van Dijk (ed.), *Pragmatics of language and literature*. *John Benjamins*. Amsterdam–Philadelphia, 141-160.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge Textbooks in Linguistics: Cambridge, UK.

- Lloyd, H. (2015). *Visual language workshop for teachers*. Museum of New Zealand Te Papa Tongarewa. <https://blog.tepapa.govt.nz/2015/04/22/visual-language-workshop-for-teachers/>
- Locke, J. (1689). An Essay Concerning Human Understanding. In: Ott, W. (Ed.), *Modern Philosophy*. BCcampus Open Publishing.
<https://opentextbc.ca/modernphilosophy/chapter/john-lockes-1632-1704-essay-concerning-human-understanding-1689/>
- Lyttle, Z (2015). *Sequential art in the Seventeenth Century: an analysis of Wenceslaus Hollar's Etchings of Genesis 12-14* (pp. 87-88). MTh(R) thesis. University of Glasgow. <https://theses.gla.ac.uk/5980/>
- MacPhee, J. (2009). *Paper politics: Socially engaged printmaking today* (with essays by D. Caplow & E. Triantafillou). PM Press.
<https://books.google.com.np/books?hl=en&lr=&id=D7dHEAAAQBAJ>
- Maiti, A. (2015). *11 Different Kinds of Ghosts We Find In Bengali Culture*. Storypick.
<https://www.storypick.com/ghosts-of-bengal/>
- Mameli, M., & Bateson, P. (2011). An evaluation of the concept of innateness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1563), 436-443.
<https://doi.org/10.1098/rstb.2010.0174>
- Maruzen-Yushodo Co. Ltd. (2004). *Choju-jinbutsu-giga facsimile edition*. Maruzen-Yushodo Co. Ltd. <https://www.facsimilefinder.com/facsimiles/choju-jinbutsu-giga-facsimile>
- Matthews, L. C., & Thakkar, B. (2012). The Impact of Globalization on Cross-Cultural Communication. *InTech*. DOI: [10.5772/45816](https://doi.org/10.5772/45816)

- Matthiessen, C. M. I. M., & Halliday, M. A. K. (1997). *Systemic functional grammar: A first step into the theory*. Retrieved from <https://www.researchgate.net/profile/Christian->
- McCloud, S. (1993). *Understanding Comics: The Invisible Art*. Kitchen Sink Press: New York, NY.
- McCurry, J. (2008). *Japan Enlists Cartoon Cat as Ambassador*.
<https://www.theguardian.com/world/2008/mar/20/japan>
- Meteyard, L. (2009). The stuff of thought: Language as a window into human nature, by Steven Pinker [Review of the book *The stuff of thought: Language as a window into human nature*, by S. Pinker]. *Opticon*1826. [DOI:10.5334/opt.070912](https://doi.org/10.5334/opt.070912)
- Mirzoeff, N. (1999). *An Introduction to Visual Culture*. London: *Routledge*. Osmond, P. 2002. [DOI: 10.4324/9780429280238](https://doi.org/10.4324/9780429280238)
- Mostowlansky, T., & Rota, A. (2020/2023). Emic and etic. In F. Stein (Ed.), *The Open Encyclopedia of Anthropology* (Facsimile of the first edition in *The Cambridge Encyclopedia of Anthropology*). <http://doi.org/10.29164/20emicetic>
- Nair, N. (2011). Cricket obsession in India: Through the lens of identity theory. *Sport in Society*. <https://www.researchgate.net/publication/254230276>
- Nasim, M. A. (2019). Disappearance of traditional games by the imitation of colonial culture through the historical parameters of cultural colonialism. *European Academic Research*, 7(11), 2777–2792. [DOI: 10.14486/IntJSCS.2019.577](https://doi.org/10.14486/IntJSCS.2019.577)
- Neuendorf, K. A. (2017). *The Content Analysis Guidebook*. Sage: Cleveland State University, USA.

- Nielsen, J. A., Zielinski, B. A., Ferguson, M. A., Lainhart, J. E., & Anderson, J. S. (2013). An evaluation of the left-brain vs. right-brain hypothesis with resting state functional connectivity magnetic resonance imaging. *PLoS One*.
<https://doi.org/10.1371/journal.pone.0071275>
- Nordquist, R. (2020). Defining Synchronic Linguistics. *ThoughtCo*.
<https://www.thoughtco.com/synchronic-linguistics-1692015>
- Nordquist, R. (2020). Definition and Examples of Diachronic Linguistics. *ThoughtCo*.
<https://www.thoughtco.com/diachronic-linguistics-term-1690385>
- Ohmann, R. (1971). Speech Acts and the Definition of Literature. *Philosophy & Rhetoric*, 4(1) 1-19.
- Peralta, E. (2020). *The differences between manga, manhwa & manhua, explained*. CBR. <https://www.cbr.com/the-differences-between-manga-manhwa-manhua-explained/>
- Pinker, S. (1994). *The language instinct: The new science of language and mind*. London-New York: Penguin Books.
- Pinker, S. (1997). *How the mind works*. W. W. Norton & Company.
- Plato. (1943). *The Republic* (B. Jowett, Trans.). Random House. (Original work published ca. 380 B.C.E.)
- Plauen, E. O. (2003). The captivating book. In *Vater und Sohn*. Ravensburger Buchverlag.
- Plauen, E. O. (2017). *The last apple*. Best Selling Picture Book.

- Postema, B. (2013). *Narrative Structure in Comics: Making Sense of Fragments (Comics Studies Monograph Series, Vol. 1)*. University Press of Mississippi.
- Pratt, H. J. (2009). Narrative in Comics. *The Journal of Aesthetics and Art Criticism*, 67(1), 107-117.
- Propp, V. (2009). *On Comic and Laughter*. University of Toronto Press, Scholarly Publishing Division. [ISBN: 9781442697201](https://doi.org/10.1017/9781442697201)
- Sakaue, Y. (2018). The Historical Creation of Kendo's Self-Image from 1895 to 1942: A Critical Analysis of an Invented Tradition. *Martial Arts Studies*.
<https://mas.cardiffuniversitypress.org/articles/10.18573/mas.66>
- Samejima, T. (2019). *A beginner's guide to Japanese art: How to read and handle emaki*. Japan Art and Culture. <https://tsumugu.yomiuri.co.jp/en/feature/a-beginners-guide-to-japanese-art-how-to-read-and-handle-emaki/>
- Sarkar, T., Salauddin, M., Paul, A., & Choudhury, T. (2023). The essence of Bengal's ethnic sweetmeats: An exploratory journey through history, tradition, and culture. *Journal of Food Quality*, 2023(12), 5008420. <https://doi.org/10.1155/2023/5008420>
- Scovel, T. (1998). *Psycholinguistics*. Oxford University Press.
- Searle, J. R. (1969). *Speech Acts*. Cambridge: Cambridge University Press.
- Sen, C. T. (1997). The Portuguese influence on Bengali cuisine. In H. Walker (Ed.), *Food on the move: Proceedings of the Oxford Symposium on Food and Cookery 1996* (pp. 288–298). Prospect Books.

- Shiko, M. (1936). *Mountain God (Yamagami)*, from the series “*Pantheon of the Gandavyuha Sutra (Kegon-fu)*”. Art Institute Chicago. <https://www.artic.edu/artworks/11249/mountain-god-yamagami-from-the-series-pantheon-of-the-gandavyuha-sutra-kegon-fu>
- Shome-Ray, A. (2013). ‘*Handa Bhonda*’ creator Narayan Debnath gets Sahitya Akademi Award: A look at the life and works of the comics writer. Lifestyle. <https://www.dnaindia.com/lifestyle/report-handa-bhonda-creator-narayan-debnath-gets-sahitya-akademi-award-a-look-at-the-life-and-works-of-the-comics-writer-1884322>
- Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts. <https://doi.org/10.1037/11256-000>
- Smith, L. T. (1999). *Decolonizing methodologies: Research and Indigenous peoples*. Zed Books; University of Otago Press.
- Spradley, J. P. (1980). *Participant Observation*. Wadsworth, Cengage Learning. New York.
- Sunder, K. (2020). *Of Mishti, Mughals and Noodles: A History of Kolkata's Street Food*. The Times of India. <https://www.forbesindia.com/article/forbes-lives/of-mishti-mughals-and-noodles-a-history-of-kolkatas-street-food/57195/1#:~:text=Mughlai%20food%2C%20from%20which%20the,and%20continue%20to%20remain%20so>
- Swartz, W. K. (2004). *Global Maps of the Growth of Japanese Marine Fisheries and Fish Consumption* (Master's thesis, The University of British Columbia) file:///C:/Users/chatt/Downloads/ubc_2004-0650.pdf
- Sweet, H. (1892). *A new English grammar, logical and historical (Vol. 1)*. Oxford University Press.

- Tambiah, S. J. (1968). The magical power of words. *Man* (pp. 175-208).
- Tang, A. (2016). *Narration and Speech and Thought Presentation in Comics* (Doctoral dissertation, University of Huddersfield).
- Theofanidis, D., & Fountouki, A. (2018). Limitations and Delimitations in the Research Process. *Perioperative Nursing*, 7(3), 155. <https://doi.org/10.5281/zenodo.2552022>
- Thiher, A. (1984). *Words in Reflection: Modern Language Theory and Postmodern Fiction*. The University of Chicago. United States of America.
- Trajan's Column. (2018). *Visual Arts Cork*. <https://www.visual-arts-cork.com>
- Turner, T. S. (1977). Narrative structure and mythopoesis: A critique and reformulation of structuralist concepts of myth, narrative and poetics. *Arethusa*, 10(1), 103–163. <https://www.jstor.org/stable/26307828>
- Veerasarn, C. (2018). Communication and perception of visual language in comic books, Khai Hua Roh: Smile of the King and Maha Sanook: When the Prince became King. *The Asian Conference on Arts & Humanities 2018 Official Conference Proceedings*. https://papers.iafor.org/wp-content/uploads/papers/acah2018/ACAH2018_39066.pdf
- Vergara, V. (2021). *Manhwa vs manga: What's the difference?* BOOK RIOT. <https://bookriot.com/manhwa-vs-manga/>
- Villines, Z. (2023). *What does the frontal lobe do?*. Medical News Today. <https://www.medicalnewstoday.com/articles/318139>
- Watson, J. B. (1913). Image and affection in behavior. *The Journal of Philosophy, Psychology and Scientific Methods*. 10(16), 421-428.

- White, M. D., & Marsh, E. E. (2006). Content analysis: A flexible methodology. *Library Trends*, 55(1), 22–45. <https://doi.org/10.1353/lib.2006.0053>
- Whitehead, T. L. (2005). *Basic classical ethnographic research methods: Secondary data analysis, fieldwork, observation/participant observation, and informal and semi-structured interviewing* (EICCARS Working Paper Series). TL Whitehead Associates.
- Widewalls Magazine (2016). Japanese Manga: An Overview of its History. *Widewalls*. <https://www.widewalls.ch/magazine/japanese-manga-comics-history>
- Wierzbicka, A. (2006). The concept of ‘dialogue’ in crosslinguistic and cross-cultural perspective. *Discourse Studies*, 8(5), 675–703. <https://doi.org/10.1177/1461445606067334>
- Willmann, A. (2012). *Storytelling in Japanese art. Metropolitan Museum of Art*. Yale University Press. https://artistproject.metmuseum.org/toah/hd/jilh/hd_jilh.htm
- Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Ed.). Wiley-Blackwell. [DOI.10.2307/2217461](https://doi.org/10.2307/2217461)
- Wolk, D. (2007). *Reading Comics: How Graphic Novels Work and What They Mean*. Da Capo Press. U. S.
- Wright, C. (1992). *Truth and objectivity*. Harvard University Press.
- Yan, J. (2019). *A Peek into Chinese Comics (Manhua)*. MR Comics. <https://medium.com/mrcomics/a-peek-into-chinese-comics-manhua-efa6a6d911b9>
- Zahin, A. R., & Nion, F. M. N. (2023). Bengali Cultural Identity in Post-Colonial Era: An Analysis of Bengali Cultural Representations. *International Journal of Research Publication and Reviews*. <https://doi.org/10.55248/gengpi.4.823.50050>