

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Narrating the Self: Multimodality and Narrative Stance in
Cartoon-Diaries

verfasst von | submitted by
Marina Grubner BA BEd

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Education (MEd)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 199 506 507 02

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Lehramt Sek (AB) Unterrichtsfach
Deutsch Unterrichtsfach Englisch

Betreut von | Supervisor:

Univ.-Prof. Mag. Dr. Alexandra Ganser-Blumenau

Acknowledgements

This long-term project would not have been finally completed without the help and support of the following people, all of whom have my heartfelt gratitude:

Thank you to **my supervisor Univ.-Prof. Mag. Dr. Alexandra Ganser-Blumenau** for constructively criticizing where the thesis desperately needed constructive criticism, and for empathically encouraging me when I desperately needed encouragement.

Thank you to **my family**. I owe you everything and will be forever grateful.

Thank you to **my friends** from all the places I have called home. Every single one of you has enriched my life.

And lastly, thank you to **my boyfriend**. My life is better because of you.

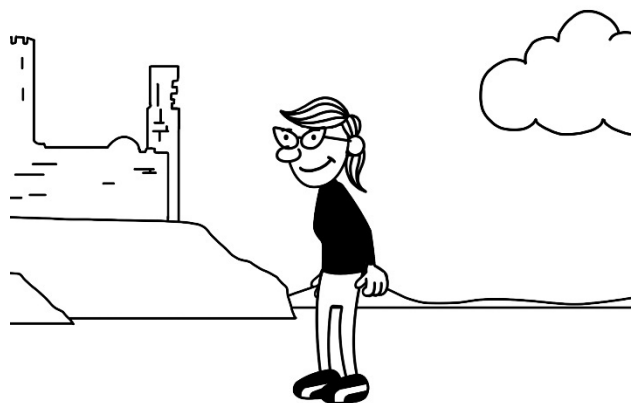


Figure 1: "Wimpy Me" created with <https://wimpykid.com/wimpyyourself/> on Jan. 3rd

Table of Contents

Acknowledgements.....	i
List of Illustrations.....	iv
1 Introduction	1
2 Multimodality.....	7
2.1 Mode – A Definition.....	9
2.2 Modal Foundations of Language and Images	14
2.2.1 Verbal Written Language	14
2.2.2 Two-Dimensional Still Imagery	16
2.2.3 Text-Image Relations.....	18
3 Diary Fiction – Genre and Narrative Structure	21
3.1 A Fiction of Factuality and Intimacy	21
3.2 Narrative Stance in Diary Novels.....	23
3.3 Diarists – Narrators and Protagonists in Diary Novels	26
4 Cartoons – Narrative Implications of a Multimodal Genre	28
4.1 Iconicity of Cartoony Imagery.....	30
4.2 Narrativity of Cartoons.....	32
4.3 Multimodal Characters	36
5 Narrativity of Multimodality – Cartoon-Diaries	39
5.1 Genre as Framework: The Diary Novel in a Multimodal Context.....	39
5.2 Temporal Structures in Multimodal Diary Narratives.....	41
5.3 The Multimodal Diarist.....	43
6 Analysis of Multimodal Narration in Contemporary Diary Novels.....	46
6.1 The Fiction of Factuality and Intimacy in Multimodal Diary Narration.....	47
6.2 Case Study I: <i>Diary of a Wimpy Kid</i>	50
6.2.1 Greg as Diarist.....	50

6.2.2 Multimodal Meaning-Making	52
6.2.3 Constructing Narrative Stance	62
6.2.4 Findings.....	68
6.3 Case Study II: <i>The Absolutely True Diary of a Part-Time Indian</i>	70
6.3.1 Junior as Diarist.....	70
6.3.2 Multimodal Meaning-Making	74
6.3.3 Constructing Narrative Stance	89
6.3.4 Findings.....	94
6.4 Comparative Discussion	95
7 Conclusion	100
Bibliography	105
Primary Sources	105
Secondary Sources.....	105
Cartoon-Diaries – Further Reading	112
Appendix.....	113
AI Acknowledgment	113
Abstract	114
Abstract [Deutsch].....	115

List of Illustrations

Figure 1: "Wimpy Me" created with https://wimpykid.com/wimpyyourself/ on Jan. 3 rd .i	
Figure 2: Greg shown in the process of journaling (<i>DWK</i> 4).....	51
Figure 3: Comparison of Greg's expression in <i>DWK</i> 27 and 216	53
Figure 4: Stilted fonts in <i>DWK</i> 42, 139, 178, 170	55
Figure 5: Verbal Written Language in the Cartoons (<i>DWK</i> 47, 63, 114).....	56
Figure 6: Onomatopoeia in <i>DWK</i> (17, 25, 74, 83)	56
Figure 7: Sound quality in <i>DWK</i> (37, 103, 18, 148, 111)	56
Figure 8: Representations of sound (<i>DWK</i> 36, 187).....	57
Figure 9: Long shot (<i>DWK</i> 56), medium-long shot (42) and mini-comic (168).....	58
Figure 10: Emotional nuance (<i>DWK</i> 5) and indirect characterization (16)	58
Figure 11: Semantic range of the cloud as visual metaphor (<i>DWK</i> 44, 106, 77, 76) ..	60
Figure 12: Showing what is not there (<i>DWK</i> 213)	60
Figure 13: Multiscenic cartoon (<i>DWK</i> 153).....	60
Figure 14: Metaphorical expressions of emotions (<i>DWK</i> 35, 73, 204)	61
Figure 15: Different levels of anger (<i>DWK</i> 113)	61
Figure 16: Outdated technology in <i>DWK</i> (23, 91, 108, 60)	62
Figure 17: Revealing emotional expressions in the cartoons (<i>DWK</i> 140, 184, 57)	63
Figure 18: Greg's attempt at reconciliation (<i>DWK</i> 189).....	64
Figure 19: Focalization shifts (<i>DWK</i> 106, 194)	67
Figure 20: Subjectivity in <i>DWK</i> (108, 41, 181, 87).....	67
Figure 21: Visual parallelism and contrast in <i>DWK</i> (32, 36).....	68
Figure 22: Sketch of Rowdy (<i>ATD</i> 23)	73
Figure 23: Different fonts in <i>ATD</i> (61, 133, 177, 99)	76
Figure 24: Labelling in the cartoons (<i>ATD</i> 124)	77
Figure 25: Representation of sound in <i>ATD</i> (137, 31, 88)	77
Figure 26: Laughter in <i>ATD</i> (204)	78
Figure 27: Varying levels of iconicity across cartoons (<i>ATD</i> 12, 174).....	80
Figure 28: Varying levels of iconicity within one cartoon (<i>ATD</i> 63)	80
Figure 29: Contrasting and parallel intersemiotic relations in <i>ATD</i> (88 and 128; 170 and 213).....	81

Figure 30: Visual metaphors in <i>ATD</i> (31, 142, 170, 112, 166, 109)	81
Figure 31: Example of page layout (<i>ATD</i> 90-91)	82
Figure 32: Direct speech in the cartoons (<i>ATD</i> 96)	83
Figure 33: Six-panel comic strips in <i>ATD</i> (120)	83
Figure 34: Visual mocking in <i>ATD</i> (171, 138, 216)	84
Figure 35: Mocking people (<i>ATD</i> 162, 5).....	85
Figure 36: Visual comparisons (<i>ATD</i> 19, 85)	85
Figure 37: Caricature (<i>ATD</i> 38)	85
Figure 38: Portraits in <i>ATD</i> (69, 71, 29)	86
Figure 39: Visualizing emotions in <i>ATD</i> (160, 53)	87
Figure 40: Reflecting cultural stereotypes (<i>ATD</i> 57)	88
Figure 41: Self-deprecation (<i>ATD</i> 121)	91
Figure 42: Book covers (<i>ATD</i> 91, 108, 213)	92
Figure 43: Split sense of self (<i>ATD</i> 182).....	93

1 Introduction

In recent years, the increasing presence of visual elements in narrative fiction has given rise to hybrid forms such as cartoon-diaries, which challenge traditional understandings of how stories are told. Situated at the intersection of diary fiction and comics, these texts combine verbal and visual modes in ways that reshape established narratological categories. While diary novels are typically characterized by a single autodiegetic voice, internal focalization, and a strong emphasis on subjectivity and immediacy, the integration of images introduces additional layers of meaning-making that complicate these conventions. As a result, questions arise as to how narrative voice, focalization, and subjectivity are constructed when storytelling is distributed across multiple modes. This thesis therefore investigates how verbal and visual elements interact to shape narrative stance in multimodal diary novels, with particular focus on *Diary of a Wimpy Kid* by Jeff Kinney and *The Absolutely True Diary of a Part-Time Indian* by Sherman Alexie. By combining insights from narratology and multimodality studies, the goal is to demonstrate how these hybrid texts both draw on and transform the conventions of diary narration.

The initial interest in the genre of cartoon-diaries emerged from the narratological implications of multimodality. The diary novel conventionally relies on first-person narration and autodiegetic focalization, presenting events exclusively through the diarist's subjective perspective. The integration of cartoons into this format, however, introduces the possibility of visually externalizing the narrating self. Although this external perspective does not constitute an objective external viewpoint, it nonetheless disrupts the strict internal focalization typically associated with diary fiction. In doing so, cartoons do not merely add illustrations within the genre, but reshape its narratological foundations by complicating voice, focalization, and subjectivity. This is why the genre of cartoon-diaries is a particularly productive site for examining how multimodal storytelling transforms established narrative conventions.

To explore these dynamics, this thesis draws primarily on the methodological framework of multimodality analyses. Additionally, theory on comics has been

consulted as well, as it offers nuanced models for understanding the interaction of visual and verbal modes within a single medium, while addressing the iconic style of the visual mode. Crucially, the cartoons in the selected diary novels do not function as decorative illustrations or simple repetitions of verbal content. Instead, they assume independent narrative functions (Eisner 165). It is precisely this dynamic interplay that makes cartoon-diaries a compelling object of study.

To examine the relationship of the visual and the verbal within the genre, this thesis conducts a qualitative multimodal analysis. Multimodality, as a broad field, encompasses virtually all forms of communication, reflecting the reality that most interactions today rely on multiple forms of expression. Its application to literature, and particularly to children's and young adult literature, is therefore both timely and relevant. In the context of cartoon-diaries, meaning is constructed through the interaction of verbal and visual modes, making a multimodal approach essential for understanding how these hybrid texts function. Such an approach also requires a broad conception of both text and reader. Therefore, a text can be any communicative artifact, including images, written language, or their combination, while a reader can be anyone who engages with the text, whether as a traditional reader, viewer, or reader-viewer, as in the case of comics.

Another reason for choosing a multimodal approach in this thesis is the fact that no methodologically similar study has been conducted before addressing the works under analysis here. While there is a vast amount of research on *Diary of a Wimpy Kid* and *The Absolutely True Diary of a Part-Time Indian*, there is a notable gap in the study of how multimodal elements shape narration in these autodiegetic texts (see the literature review below for an overview of related studies). This thesis aims to address exactly this gap. Building on the premise that multimodality influences narratological structures, the central research question guiding this investigation is:

In what ways do visual and verbal modes interact to construct narrative stance in cartoon-diaries such as *Diary of a Wimpy Kid* and *The Absolutely True Diary of a Part-Time Indian*?

Here, "narrative stance" refers to the combined effect of the narrator's voice (i.e. who tells the story) and the narrative focalization (i.e. who perceives the story), as well as

the narrator's subjective positioning. These narratological categories – voice, focalization, and subjectivity – thus play a central role in the analysis and will be investigated as to how they are shaped and complicated by the interplay of verbal and visual modes.

Before discussing the theoretical and methodological framework that underpins the analysis in the following chapters, a brief discussion of the two selected works will be provided. On the one hand, their selection as examples of multimodal diary fiction shall be justified, and on the other hand, a concise literature review shall highlight the novelty and relevance of applying a multimodal approach to these novels. From now on, the abbreviations *DWK* (= *Diary of a Wimpy Kid*) and *ATD* (= *The Absolutely True Diary of a Part-Time Indian*) will be used for the sake of brevity and readability.

It is almost impossible to ignore the immense success both novels have had. They are among the most prominent examples of diary fiction for children, teenagers, and young adults, which partly motivates their selection for this study. While *DWK* and *ATD* were not the first cartoon-diaries, they are certainly among the most commercially successful.¹ Further, from a personal pedagogical perspective, the focus on children's and YA literature is particularly relevant, and to my knowledge, there are no comparable cartoon-diaries aimed at adults. While at a first glance both novels appear typical of the cartoon-diary genre, they differ in significant ways, making them well-suited for a comparative analysis. This contrast provides a solid foundation for a genre-based investigation, allowing for insights into the scope and variation within the genre.

First, *DWK* is an ongoing series created by Jeff Kinney, who is both the author and illustrator. It follows Greg Heffley's first year in middle school, as he records problems at school, with his family (his parents and two brothers), and his experiences with his best friend Rowley in his diary. To date, the series comprises 20 volumes, with the

¹ Another notable example of this genre is the *Amelia's Notebook* series by Marissa Moss, which began with its first volume in 1995 and can be considered the first of its kind. The debut title received multiple awards and enjoyed considerable popularity until the series concluded in 2015 with *Amelia's Graduation Yearbook*, meaning that, unlike in the *DWK*-series, the protagonist ages over time. For further information visit the series' website: <https://www.marissamoss.com/books/amelias-notebook/> (Moss).

most recent one published in 2025, and has sold an astonishing 300 million copies worldwide (Heritage).² Considering the overly large scope of the full series, this thesis focuses solely on the first volume published in 2007, as a representative text for analysis. The series' remarkable popularity has been attributed partly to its reflection of contemporary childhood experiences (Voinea and Norel), but also to its multimodal format, which may lower barriers for reluctant or struggling readers due to the inclusion of imagery (Voinea and Norel; Arend and Mahrt).³ As boys in particular are often stereotypically associated with reading reluctance, *DWK* is often cited as a text that successfully motivates young male readers (e.g. Gleason). Additionally, the series' distinctive humor has been identified as a further key factor for its success (e.g. Mroczek-Bayci).

Although the role of multimodality in *DWK* has been discussed (e.g. Reid and Serafini; Rickard; Verusha Ali et al.), multimodal implications on the novel's narratology remain largely unexplored. In turn, narrative voice in *DWK* has been examined, but without a detailed consideration of the text's multimodal nature (Hui, "Narrative Voice"). Furthermore, no genre-based studies have been done to situate the novel within an established practice of meaning making.

One area of *DWK* that has received considerable attention is its multimodality in the context of translation studies. Researchers have explored the challenges of translating a multimodal text, particularly in relation to expressive language (such as onomatopoeia), ideophones, metaphors, or character names, into languages such as German, Croatian, Serbian (Prodanović Stankić and Begonja), Italian (Pischedda), Indonesian (Arfianti et al.), and German and Slovenian (Haller). Translation studies become especially noteworthy when ideological motivations lead to alterations in both the verbal and visual elements of the Arabic version (Alós; Al-Jabri).

² In 2014, the series even earned the Guinness World Record for "most copies of a fictional diary sold." At that time the series encompassed 9 volumes and had sold over 150 million copies ("Most Printed Fictional Diary Series").

³ Although the series has been marketed as an easy read due to the presence of imagery, Rickard argues that the books are popular not because of their simplicity, but because they in fact "mask complexity" (54).

Further research on *DWK* has examined a range of aspects, including power relations between adults and children (Kurnia et al.), constructions of masculine hegemony (Potter; Aktaş Kubal and Çetinkaya), the fetishization of money (Hui, “Currency”), or Greg’s transition from childhood to adolescence from a psychosocial perspective (Santoso and Fajar).

Next, *ATD* is a semi-autobiographical novel by Sherman Alexie, published in 2007 and illustrated by Ellen Forney. The protagonist, Arnold Spirit Jr., is a Spokane teenager living on a reservation who narrates his experiences of transferring to an all-white school in a nearby town, as well as his struggles within his family and his friendship with his best friend Rowdy. Like *DWK*, the novel achieved considerable commercial and critical success, appearing on several “Best Books of the Year” lists in 2007, including that of “The New York Times” (Suico). Similar to *DWK*, the novel’s illustrated format has sometimes led to the assumption that it constitutes an easy read, however, scholars argue that its popularity stems not from its simplicity but aesthetic complexity (MacIntyre) and its strategic use of humor, which has also been noted as particularly effective in engaging young male readers (Elstad).

With regard to the novel’s multimodality, research has typically focused on Junior’s motivation for drawing. However, to my knowledge, a systematic narratological analysis of *ATD*’s multimodal structures has not yet been undertaken. Still, many scholars mention the function of Junior’s drawings as a means to navigate his identity crisis, which has been described as a “split” identity (Garić). However, it is warned against reading Junior as merely torn between two static cultures, but instead framing his character within a dynamic and political nature of cultural identity (Suhr-Sytsma).

Beyond multimodality, research has predominantly situated the novel within Native American and YA literature studies. Scholars read the text as a typical YA novel in structural terms, yet one that is deeply shaped by its specific cultural context, portraying the protagonist’s negotiation of complex identities across cultural and colonial boundaries (Sparks; Parsegova; Bodden). Issues such as assimilative schooling pressures faced by Native Americans (Katanski), concepts of survivance under colonial oppression (Doğru; Johnson), and alcoholism on reservations (Lacy) have been central to critical discussions. Despite the dark and partly traumatic

themes in the novel, many researchers conclude that it ultimately conveys an empowering and hopeful message, in line with conventions often associated with YA literature (Unterweger; Palmer; Vogel). At the same time, some critics argue that, given its predominantly white readership, the novel risks reinforcing stereotypes about Native Americans (Wahpeconiah; Szeghi), while others advocate for its careful and context-sensitive use in classrooms, particularly with students of similar cultural backgrounds (Spring). Within a YA context, Junior's characterization has also been examined through the lens of disability studies, as his hydrocephalus and concomitant conditions mark him as different (McIntyre and Brown), which in turn has been interpreted in relation to broader cultural markedness (Vogel; Wahpeconiah).

Much like for *DWK*, many scholars highlight the novel's humor and see it as a central coping mechanism that enables both protagonist and readers to navigate the novel's darker themes (e.g. Lacy). Further studies, which also foreground the Native American context, have addressed the translation of taboos (Nazari Robati and Zand), critiques of normative happiness scripts (Wayland), the symbolic role of basketball (Goldstein), or constructions of nomadic identity (Moon). Works that deal with pedagogical approaches to teaching the novel will be discussed in the conclusion of this thesis.

To sum up this short literature review, both novels are highly popular yet provide different perspectives within the genre of cartoon-diaries. They target different audiences from children to young adults and explore varied themes, from everyday school life to Native American experiences. These differences make a comparative analysis especially valuable. Prior to this, the next chapters will lay out the theoretical and methodological foundations for the study.

2 Multimodality

Today, multimodality is everywhere, and we are constantly surrounded by multimodal environments, whether in our daily lives, which are dominated by screens, in our free time on social media or when working in a digital environment. However, this is not a recent development but has been evolving for decades, as early research in multimodality already argued that there are no purely monomodal texts (Mitchell, “Media”) and that every text has multimodal potential (Kress). The evolution of media makes the study of multimodality tricky, since it is an endeavor under ever-changing conditions, as are the modes and the ways that people use them (Kress and van Leeuwen, *Multimodal Discourse*; Bateman et al., *Multimodality*). When Kress and van Leeuwen came up with their concept of multimodality in 2001, the media landscape certainly was different than it is today, 25 years later. The authors define multimodality as “the use of several semantic modes in the design of a semiotic product or event” (*Multimodal Discourse* 20).⁴ The essence of their theory is that different modes are combined on a ‘design’ level and are then realized on a material production level. This process also relates to intermediality and often encompasses different media in the production process. The production as well as interpretation process – including the producers and interpreters – play a crucial role in Kress and van Leeuwen’s theory as well. Although communication requires both articulation and interpretation to take place, its success heavily depends on the latter, as the interpreters are expected to have semiotic knowledge while simultaneously being subjected to a certain discourse, i.e. certain positions in society and concomitant values. To what extent the producer’s intention and the interpreter’s interpretation match is further determined by the context of the reception (*Multimodal Discourse* 8).⁵ The social semiotic aspect of multimodality has since been highlighted by many researchers and the focus of this scientific field certainly lies on the production of meaning (Wildfeuer et al.). At its core, this

⁴ The significance of semiotics within multimodality studies is already evident in this definition. Scholars such as Bateman et al. draw extensively on the foundational work of Ferdinand de Saussure and Charles Sanders Peirce, whose theories fundamentally shaped the field of semiotics. A more detailed engagement with this theoretical groundwork, however, would go beyond the scope of the present thesis and will therefore not be pursued further here.

⁵ Kress and van Leeuwen determine four levels: discourse, design, production and distribution. These dimensions provide the context for multimodal analyses (*Multimodal Discourse*).

production relies on the combination of various modes, however, this process is rather complex and exceeds mere combination. In this respect, Stöckl speaks of a highly genre-sensitive form of co-contextualization rather than combination: “[O]ne mode is contextualized by another, so that modes mutually provide co(n)text and can often only function properly through this co-contextualization” (41). This means that two (or more) modes do not merely translate or support each other, which would presume a kind of dependency or hierarchy. Rather, the modes reinforce and question each other, thus creating meaning in a constant dynamic interaction. Therefore, the core of multimodal analyses is investigating exactly this dynamic interplay, instead of a singular addition of modes (e.g. Sousanis; Kukkonen, “Transmedial Narratology”).

Naturally, this process remains largely opaque, thus constituting a central concern for researchers in this field. Lemke has famously coined the metaphor of viewing multimodal meaning-making not as the sum of adding modes, but as seeing it as a multiplication, highlighting their interrelation. Regardless of the metaphor employed, whether as additive or factorial, it becomes evident that modes are interdependent and create meaning in a dynamic interplay that cannot be fully disentangled to distinct components. One mode might have already been influenced by another before the former can be described, making it impossible to reconstruct this process (Bateman et al., *Multimodality* 16–17). Other researchers use the term *intersemiosis*, which highlights both the interaction between semiotic resources, as well as the active role of reception. Packard et al., for example, claim that multimodal meaning-making depends heavily on the recipients prior world knowledge, particularly with metaphoric, non-literal relations between modes (59–60).

Further complexities, such as societal power relations (cf. Bateman et al., *Multimodality* 113–16), or spatial and temporal gaps between the producer(s) and interpreter(s) (cf. Kress and van Leeuwen, *Multimodal Discourse* 113), highlight just how intricate multimodal communication can be. At the same time, multimodality research has been criticized for its strong grounding in semiotics and linguistics, which can result in privileging written language over other modes (Kress and van Leeuwen, *Reading Images*; Bateman et al., *Multimodality*). Frameworks such as that proposed by Bateman et al. seek to counterbalance this tendency. To examine how meaning is

constructed in multimodal artefacts, the following section therefore outlines and clarifies several key concepts in multimodality studies.

2.1 Mode – A Definition

Globally speaking, modes are comprised of semiotic resources, which are realized on a material level in a medium, e.g. writing is the mode, and its semiotic resources are words, font, layout, punctuation, etc., and it is realized in a medium such as a novel, a travel blog, etc. (cf. Stöckl 46–47; Kukkonen, “Transmedial Narratology” 35). As far as the social dimension of a mode is concerned, Kress sees it as “a culturally and socially fashioned resource for representation and communication” (45).

A more precise conception of mode requires a closer investigation of both its material realization and the semiotic resources it mobilizes. Such a differentiated yet flexible conceptualization of the semiotic mode is given by Bateman et al., who define a mode as “a constellation of practice within a community of users” (“Definitions” 318), who establish meaning based on three semiotic levels that determine every mode:

- “(i) a deformable perceptible materiality, potentially involving multiple sensory channels (*canvas*),
 - (ii) a classification (paradigmatic) of formal units and structures (syntagmatic), which defines the material deformations that are pertinent for the semiotic mode, and
 - (iii) a level of *discourse semantics* that provides dynamic mechanisms for the abductive construction of discourse structures assigning contextual interpretations to the form classifications deployed”
- (“Definitions” 318–19, original emphasis)

In order to grasp the modes – written language and iconic imagery – that are described later in this chapter in general and then with regard to their specific realization in the primary texts, the components of this definition warrant a closer explanation:

(i) Canvas and affordances

First, the canvas, a constitutive feature in Bateman et al.’s definition, is a core element in any communicative situation, apart from the sign producer(s) and consumer(s). Canvases are “the locus of semiotic activity” (*Multimodality* 101) and can be tangible, such as pieces of paper, screens, or immaterial like face-to-face interactions. In each

case, they provide the spatial and material conditions under which signs are produced, organized, and interpreted, and are thus shaped not only by physical properties but also by the social conventions governing their production and reception (*Multimodality* 111).

As the site where meaning emerges, a canvas is characterized by specific affordances,⁶ which are “the possibilities for action that it opens up for an agent in an environment” (*Multimodality* 90). A printed page in a novel, for instance, affords actions such as turning pages, arranging text spatially, or using different types of (usually only black) ink in particular ways. An object’s design can influence its affordances, which means that one and the same medium (e.g. a sheet of paper) can be designed differently and, for example, absorb more or less ink, be more or less thick, etc., which in turn has an influence on how meaning can be created. However, what a medium can afford is not limited necessarily to its material specifications, since a non-moving canvas is in fact very much able to depict movement.⁷ For the purposes of multimodal analysis, however, it is less productive to catalogue all potential affordances than to examine those that are actually realized in a given artifact, as only these specifically contribute to the processes of meaning-making under investigation (*Multimodality* 90–102).

(ii) Form

A mode’s formal organization determines how it is being used, for example the mode of sound can be used in music as well as spoken language. In both these situations, the mode is the same but is made use of differently. Therefore, talking of a visual mode can encompass drawings as well as written language, both of which clearly afford different options spatially and/or temporally. Further, a mode’s form is set on a continuum of lexical to grammatical organization, where lexically organized modes comprise a stock of signs, whereas grammatically organized modes work their resources through a productive system of finite rules to make infinite meanings. Depending on its organization, a mode can thus form paradigmatic or syntagmatic relations

⁶ Different terms than ‘affordances’ have been used by researchers. For example, Kress uses the term ‘reach’ to denote what a mode’s material carrier can afford.

⁷ See Chapter 4.2 for a short elaboration on how movement can be depicted in a still canvas.

(Bateman et al., *Multimodality* 114–16). Kress and van Leeuwen emphasize that not only is no mode strictly organized in one way or the other, but also that one mode can fulfil, for example, a paradigmatic purpose for one group of people (e.g. the producers) and simultaneously a syntagmatic one for a different group (e.g. the consumers) (cf. *Multimodal Discourse* 113).⁸

(iii) Discourse

The second level of how a mode's content is organized is the discourse level, more precisely its discourse semantics, which demarcates the mechanisms of interpreting a sign in context, generally addressing the question of how the readers actually make sense of the material in front of or around them. How meaning is created and interpreted is highly sensitive to context. Thus, instead of investigating all potential ways of meaning-making, it is more productive to analyze how a mode's affordances are actually realized in specific contexts (Bateman et al., *Multimodality* 118).⁹ Additionally, Bateman et al. argue that having fixed assumptions of what a mode can afford can be misleading. It is for example presupposed that the mode of drawing affords meaning based on spatial and simultaneous logic, whereas written language follows temporal and sequential logic. The very genre of comics is proof that although the organization might be either spatial or temporal, still, spatiality and temporality can be expressed independent of how the content is organized (see Chapters 2.2.3 and 4.2). Therefore, Bateman et al. argue that a description of a mode's affordances should be based on their perception.¹⁰ A multimodal analysis should not investigate what could

⁸ Kress and van Leeuwen mention the intriguing example of the mode of smell. Whereas for most people, smell is organized as distinct entities that each evoke different memories, for scientists, perfumers, or aromatherapists there is a grammar to it, by creating a combination from about 50 distinct smells in a syntax of the formula 'head – body – base' to produce a complex "smell-sign" (*Multimodal Discourse* 113).

⁹ The following situation should exemplify this: Although a visual depiction requires a positioning that is sensitive to location and relative positioning, the degree to which these aspects are necessarily realized differs according to the situation. An architect's plan that is supposed to be submitted to authorities for receiving a permission for building is, and as a matter of fact has to be, far more precise than his:her sketch to present to a customer to enable a broad overview of a project long before it even begins. Thus, it is more meaningful and productive to investigate the specific situation, i.e. the general sketch or the submitted plan, and analyze how the visual mode creates meaning and is interpreted, instead of analyzing what would be potentially measurable (cf. Bateman et al., *Multimodality* 93–94).

¹⁰ Perception in this case, however, does not limit itself to a specific and singular sensory perception, as modes tend to cross senses as well (Bateman et al., "Definitions" 319).

be done within a mode, but what is done. Whether meaning is then organized spatially or temporally, for example, is a matter of discourse semantics. As there are no universal shortcuts to meaning-making, differences in the properties of similar looking representations have to be explained by discourse semantics (*Multimodality* 117–19). The discourse semantics are not inherent within an individual mode, but are located at the level of text or even genre (Stöckl 48).

In summary, there are three layers of content to a semiotic mode: A mode has i) a material level (i.e. the canvas), as well as a certain ii) form (i.e. the “technical features”), and is embedded in iii) discourse semantics (Bateman et al., *Multimodality* 117). This definition makes it clear, for example, that it is not possible to talk of a language mode per se. Rather, a linguistic quality has to be further subdivided into spoken verbal and written language. Although they share the same lexicogrammar and a general temporal organization, both forms have different canvases (e.g. a monologue on stage and a page of a novel), as well as a different graphic-spatial perception (gesture, volume, prosody vs. punctuation, layout, color, size), and are also interpreted in different discursive structures. Thus, a distinction between semiotic and material affordances enables a precise differentiation of two modes (e.g. spoken and written language) of the same mode family (e.g. verbal language) (Bateman et al., *Multimodality* 122; Stöckl 48).

Apart from these three defining aspects of a mode, socially established practices through which meaning is produced and interpreted play a crucial role when engaging in multimodality studies. Within this context, the concept of media becomes particularly relevant. The notion of a medium should not be conflated with that of the canvas, which refers more narrowly to the material surface of a mode’s inscription. Rather, media provide specific communicative channels, which are themselves shaped by discourse semantics, making them a central category in multimodal research. Marshall McLuhan’s famous quote that “The medium is the message” (7) emphasizes that media are not neutral carriers of meaning, but shape how the content is perceived and understood. At present, researchers like Wolf also argue that a medium is never just a neutral vessel for transmitting (narrative) material, but significantly shapes the content (263). A meaningful definition of media is given by Bateman et al.,

who emphasize the established communicative formats of media: Therefore, a medium is “a historically stabilized site for deployment and distribution of some selection of semiotic modes for the achievement of varied communicative purposes” (*Multimodality* 123). Importantly, while a medium carries a semiotic mode, it does not necessarily have to realize its full range of affordances. For example, a telephone call cannot rely on all affordances of spoken interaction, as aspects like gestures or mimics cannot be transported. For Bateman et al., media are therefore best conceived as communicative forms rather than physical entities, which is why they consider a tablet not to be a medium but a canvas that can realize various media, such as a newspaper article. Other researchers also acknowledge the lack of clear boundaries in defining media; Doloughan, for example, argues that language itself can function both as a medium and a mode. As a medium, language can relay a story and as a mode it affords different properties shaping said narrative (6). Such perspectives underline the interdependence of modal and medial categories, showing how different media can be interconnected through their shared or overlapping use of semiotic modes. This overlap creates relationships between media that may share similar meaning potentials while differing in degrees of realism, materiality, or permanence.¹¹ Consequently, media have a decisive impact on how content is communicated, depending on their particular qualities, i.e. whether they are static or dynamic, two- or three-dimensional, transient or durable, etc. (cf. Bateman et al., *Multimodality* 104). With this understanding of what constitutes a mode, the following section will further specify the characteristics of written language and iconic imagery.

¹¹ Bateman et al. offer the example of embedding photographs in different media: If a photograph is shown in a movie, the expectations of the two modes ‘photograph’ and ‘movie’ remain unaltered (*Multimodality* 124–27). However, Packard et al. argue that showing a (drawn replication of a) photograph in a comic does change its modality, because what is shown is in fact not a typical photograph, but a representation thereof. Still, the depiction does share a meaning potential that is very closely related to photographs outside of comics (53). The question here seems to revolve around how close to the actual real-life counterpart the embedded mode is. Showing an actual photograph in a movie does not alter the original properties of the image, but just determines how it can be viewed, whereas a reconstruction of a photograph in a drawn medium does estrange it from its real-life counterpart to some extent.

2.2 Modal Foundations of Language and Images

Although multimodal meaning does not stem from a simple combination of isolated features of individual modes, but a dynamic and complex interaction, it is analytically useful to consider the affordances of the modes that participate in the genre of cartoon-diaries for the analysis in Chapter 6. In other words, isolating the modes for analytical purposes does not imply that they function independently in practice, but serves to establish a descriptive baseline against which their later interaction can be more precisely assessed. Accordingly, this chapter first outlines the fundamental characteristics and semiotic resources of the written and the cartoon mode, and then discusses how each typically generates meaning. While the primary goal is not yet to analyze an integrated multimodal genre, certain implications of their general co-occurrence will be discussed.

2.2.1 Verbal Written Language

As far as verbal written language as a mode is concerned, it can be materialized on various canvases, such as a page in a novel, a screen, a blackboard etc. The content is formally organized both lexically and grammatically, thus forming paradigmatic as well as syntagmatic relationships. Additionally, written language affords a temporal sequence (Kress 4). The discourse level, i.e. how meaning is made and interpreted by the reader based on the signs used, has manifold configurations and there is a vast amount of different genres that this mode participates in.

What separates language from other forms of meaning making are certain properties, which Bateman et al. identify as conventionality, compositionality, substitutivity, and functionality. Although these properties are also shared by other modes in different ways, they are crucial for defining the term ‘language’ (*Multimodality* 25). First, language is conventional in that the relation of its signs and their real-life counterparts is arbitrary and relies on traditions of using them. There is, for example, no close connection between an actual dog and the word itself, neither in written form nor as a sequence of in themselves meaningless sounds.¹² Choosing one out of several,

¹² It could be argued that both words and images share a common origin, since letters historically evolved from stylized images, as apparent in Egyptian hieroglyphs, for example (cf. Eisner 8–9). Today,

potentially equally successful, options, which is then used repeatedly by a community is essential for the existence of language (Lewis, qtd. in Bateman et al., *Multimodality* 38). For Kress, the arbitrary relationship of word and object means that words are inherently meaningless in that their arbitrary sound- or letter-shape has to be filled with meaning by the reader (3).¹³ Further, language is constituted by its compositionality, i.e. the fact that the meaning of a complex entity can be understood by deciphering the meanings of its parts. Closely related to this aspect of compositionality is substitutivity, meaning that one part of the complex chain can be substituted by another (e.g. changing the subject of a sentence) without it affecting the meaning of the other parts, thus forming a paradigmatic relationship. Moreover, the parts of language are inherently structured according to different functions they fulfil, inscribing a functionality aspect in language, which mainly refers to communicative functions of systemic-functional linguistics (i.e. ideational, interpersonal, and textual). The communicative purpose of language is simultaneously practiced within society as well as shaped by it (Bateman et al., *Multimodality* 49).

While these are properties of language in general, as mentioned before, one mode family can have various modes with distinct affordances. To mention one example, Kress contrasts written and spoken language and argues that e.g. framing works differently in both modes: Written language uses spacing and punctuation, whereas spoken language relies on pausing and intonation (127). Despite these differences, written and spoken language are also interrelated. Kress argues that reading written language actually affords some properties from spoken language, as most readers tend to subvocalize (151–52).

Reflecting on the concept of framing in verbal language demonstrates that the affordances of different modes overlap, since framing is not limited to the visual arts only (cf. Kress and van Leeuwen, *Multimodal Discourse* 2–3), but also operates within written language. In this sense, written language can itself be understood as a visual

however, the connection between symbols and letters has become so distant that the modern alphabet must be seen as an arbitrary system for meaning making.

¹³ A relatable example thereof is when readers enjoy a book, but are utterly disappointed upon watching the screen adaptation, which forces them to see a version of how somebody else filled the meaningless word vehicles with his:her imagination (Kress 3).

mode, particularly when considering the significance of layout. Elements such as graphic design, spatial positioning on the page, and the semiotic dimensions of what is expressed all contribute to meaning-making (Packard et al. 34). Accordingly, Kress and van Leeuwen claim that “writing is itself a form of visual communication” (*Reading Images* 17). Nevertheless, when the term ‘visual mode’ is used in this thesis, it primarily refers to pictorial elements, a distinction that will be clarified in the following subsection.

2.2.2 Two-Dimensional Still Imagery

Bateman et al. describe visuality as the meaning-making processes that follow the effects of light on the eyes (*Multimodality* 30). Thus, describing a singular visual mode is not possible, as that would encompass all communicative artefacts that are perceived through an interpreter’s eyes. As mentioned above, also written language is a visual mode and, given its materiality, relies on visually perceivable affordances such as layout, spacing, or sizing. Despite the fuzziness of the term ‘visual mode,’ this thesis employs it in a more specific sense to denote the ‘pictorial mode.’ The focus thus lies not merely on visual perception in general, but on the particular semiotic resources associated with imagery. Generally, pictorial refers to all forms of visual representation such as drawings, paintings, photographs, or other static images,¹⁴ and of course encompasses cartoons as well. At this point, a further distinction must be made between cartoons – or cartoony imagery – as a mode and cartoons as a medium. Cartoony imagery as a mode consists of iconic drawings, whereas cartoons as a medium constitute multimodal artefacts in their own right, combining written language with iconic images. For the purposes of the analysis in Chapter 6, therefore, speaking of a visual or pictorial mode, while acknowledging the broad conception of

¹⁴ According to an online dictionary, the term “image” also encompasses statues, so a pictorial mode could also be realized as a three dimensional artefact (Oxford Learner’s Dictionary). This illustrates how broadly a visual as well as pictorial mode can be conceptualized and highlights the need for context-dependent definition.

the term, focuses on the comic-specific iconic drawings characteristic of cartoons,¹⁵ so as not to exceed the scope of this thesis.

The fundamental characteristics of cartoon-specific iconic imagery are shared by other forms of pictorial modes as well, and will be described briefly, in line with Bateman et al.'s definition as in the previous subchapter. Similar to written language, the pictorial mode can be realized on a variety of canvases, ranging from the printed page to digital screens or even walls, depending on the material and technological context. Instead of affording temporal sequence, images usually make meaning through spatial organization and are perceived instantaneously (Kress and van Leeuwen, *Reading Images*). Given the wide range of possible image configurations, the discourse level of this mode is likewise highly diverse. Compared to language, the content of an image is organized more lexically than grammatically.¹⁶ General properties defining pictorial media are that depictions often work based on associativity, non-compositionality, resemblance, non-substitutivity, and functionality. It is again Bateman et al. who describe these functions:

Associations play a crucial role when decoding visuals, especially compared to language. Particularly in the case of iconic depictions, the human visual processing system is highly effective and productive (*Multimodality* 33). Moreover, as opposed to linguistic decoding, visual perception is not compositional, as a form (e.g. a cat-shape) is not recognized by perceiving the individual parts (i.e. a circle as the head and two triangles as the ears) and combining them to a whole, but instead the whole is recognized instantaneously. Only then, assumptions about the individual parts can be made as a top-down process (*Multimodality* 33). This is also closely related to processes of resemblance. Although images are more readily associated with

¹⁵ With respect to comics specifically, Packard et al., for example, speak of “[comicspezifische] Zeichenmodalitäten,” denoting the iconic semiotic resources employed in the medium of comics (cf. 52 pp.).

¹⁶ In “Reading Images: The Grammar of Visual Design,” Kress and van Leeuwen do propose a form of visual grammar, despite frequent criticism that multimodality relies too heavily on linguistic models. They deliberately adopt the term ‘grammar’ to establish common ground between two of the most prominent semiotic modes, written language and pictorial depictions. However, their use of the term does not suggest a fixed system of rules comparable to traditional linguistic grammar. Rather, it refers to socially and culturally shaped patterns of meaning-making and recurring regularities. For instance, conventions of sizing and placement may signal hierarchy: A character positioned centrally and depicted on a larger scale than others typically conveys greater importance or authority within the image.

resemblance to their real-world referents (e.g. a dog versus a drawing of a dog) than language is (e.g. the animal versus the word ‘dog’), Bateman et al. note that pictorial representations also vary in their degree of conventionality. Nevertheless, this resemblance can be so strong that children are often able to recognize a drawing of a dog before they are able to read and decode the written word, a skill that must be explicitly learned (*Multimodality* 34).¹⁷ As opposed to language, visuals are exempt from substitutivity. While it does not affect the other parts of a sentence, if one word is exchanged, changing a visual element of a picture (e.g. swapping an elephant for a giraffe in a drawing of the African veldt) does affect how the surrounding is perceived, simply because of the shape and size of the substitute (*Multimodality* 39). Bateman et al. argue that images also realize the same metafunctions as language, as these are not confined to verbal communication but extend to other semiotic modes as well (*Multimodality* 49–50).

2.2.3 Text-Image Relations

Having outlined the central properties of written language and images, this section now focuses on the implications of their interaction. Since multimodal meaning emerges from the interplay rather than an isolated combination of modes, it is necessary to examine how their distinct affordances, such as the temporal logic of writing and the spatial logic of images, overlap, and influence one another. To this end, insights from comics studies seem intriguing, as comics are a medium in which verbal and visual modes jointly construct narrative meaning.

Although both modes are visual, i.e. they are perceived through the interpreters eyes, Kress argues that the two modes are inherently differently organized and thus diverge in the way they make meaning, and proclaims that “[t]he world told is a different world to **the world shown**.” (1, original emphasis). He argues that language is able to form “genre[s] of narrative”, while images develop “genre[s] of display”, and

¹⁷ The semiotician Charles Sanders Peirce divided signs into icons, indexes, and symbols. He postulates that icons are the signs that resemble their real-life counterpart (i.e. the object) the most, as opposed to symbols, which are much more conventionalized. Still, Bateman et al. note that even conventionalized signs such as symbols have to be resembling enough for the reader to establish the relation to the real-life object (*Multimodality* 59).

concludes that “[t]he world narrated’ is different from ‘the world depicted or displayed’” (2).¹⁸

This crucial difference between the content organization of written language and images is discussed in great detail in research literature, thereof, the following sections will be based on Kress, Sousanis, and Groensteen (*System*). These researchers claim that language is innately organized temporally, which means that in a verbal utterance there is a first word, words in between, and a last word, determining a sequence in which meaning has to be produced and interpreted. Language is therefore linear and sequential. In contrast, images constitute themselves through their spatial organization, where the content expresses relations and locations. Instead of a sequential progression, the content is rhizomatic and spreads like a network or system. The interpretation process therefore cannot be sequential, but is instantaneous¹⁹ and simultaneous, where the content is taken in all at once.²⁰

However, a strict division of temporality belonging to language and spatiality belonging to visuality, has been criticized by researchers (cf. Doloughan; Bateman et al., *Multimodality*). Especially in mixed genres, such as comics or picture books, Doloughan sees a reciprocal relation of these two properties. She argues that images are able to connect over time as well, thus accomplishing a “temporalization of the spatial” (34). For example, when images are strewn across a novel, where they are assigned a fixed place and relate to events that have previously been temporalized in the verbal

¹⁸ Narratologically, this correlates to “showing vs. telling” (Genette 163). Here, Genette refers to the narrative techniques of mimesis (i.e. a narrator is not present, as in theatre or dialogue) and diegesis (a narrator makes a story accessible to an audience). He claims that there are varying degrees of mediation, so narrators can be more or less present in a story, but also discusses these terms critically (cf. 162–85). Thus, even within the mode of written language, there are genres that rather display than narrate and vice versa. The implications of multimodality on narratology will be discussed in Chapter 5.

¹⁹ The instantaneity of images potentially opens a more immediate access to the viewer compared to a verbal mode, as has been noted by Kress. Additionally, he claims that images are determinate and present a given, while language in the form of sound-letter shapes needs to be filled in and is more schematic. This activation of the readers’ imagination of language, according to Eisner, gives verbal narration a higher participatory effect, as opposed to purely visual narration (147–48). However, the process of closure, which will be described in greater detail in Chapter 4.1, shows that reader involvement is also especially high in settings that employ iconic imagery, such as comics (cf. McCloud).

²⁰ Groensteen (*System*), as well as Sousanis, argue that sequential and simultaneous reception relate to two different types of awarenesses, i.e. analytical and holistic perception. Sousanis claims that, in comics, both these perception processes are combined because of the co-presence of “sequential and simultaneous modes” (63), which is perceived in a “dynamic cycle of read-look, look-read” (64).

narrative, they also gain sequentiality. At the same time, if images are inserted in running text, they lead to a “spatialization of the temporal” (34), by shaping the spatial organization and the layout of the text. The images can draw focus and be interpreted before, after, or during the running text, thus interfering with the temporal sequence of the verbal narrative. This illustrates a key effect of multimodality: Embedding modes within each other, while still retaining their inherent structure, reshapes them, allowing images to gain a temporal quality, and words a spatial one.

This process, however, has its limitations, since it does not entail an inherent change of a mode’s fundamental organization, even though certain approximations may occur. In comics in particular, temporality and movement can be conveyed through visual metaphors (see Chapter 4.2), allowing spatial arrangements to express temporal progression (cf. McCloud; Wolf). Yet this remains a representational strategy rather than a structural alteration of the mode itself. As a further example, Sousanis refers to the stylistic device of ekphrasis (i.e. when an object or person is described in vivid detail). Although this has been cited as an example of convergence between written language and imagery (Mitchell, “Showing Seeing” 174), a verbal description of an image can never fully represent said image, but always remains “a representation of thinking about having seen a picture [...] While image is, text is always about” (Sousanis 58).²¹

Based on these considerations it can be said that neither the verbal nor the pictorial can ever substitute the other. Additionally, the differing organizational principles of images and language should not be understood hierarchically or imply the superiority of one mode over the other. While language appears inherently complex due to its reliance on grammatical structures, Kress and van Leeuwen argue for a comparable grammar of visual design, highlighting that both language and images can exhibit varying degrees of complexity (*Reading Images* 107). Moreover, a mode’s creative potential is not dependent on the absence of formal rules. Just as images can unfold

²¹ Other instances, where modes and their perceptory channels overlap, are for example, as mentioned above, the process of subvocalizing, or also when a text is read aloud. Here, written language gets an auditory component (Mitchell, “Showing Seeing”; Kress). Another overlap of modes is haptic visibility, which describes having sensory reactions to visual input. Bateman et al. here give the example of seeing someone in a movie scratch a blackboard (*Multimodality* 36–37).

unbound creative potential, literary movements such as Dadaism demonstrate that language, too, can be creative outside of grammatical norms. Therefore, verbal and pictorial modes must be understood as distinct yet equally significant meaning-making resources that engage in interplay to jointly construct meaning in ways neither mode could achieve alone.

3 Diary Fiction – Genre and Narrative Structure

After the methodological discussion in the previous section, the following two chapters provide the narratological framework for the multimodal analysis in Chapter 6. Given the fact that multimodal meaning-making processes are highly genre sensitive (cf. Stöckl), it is crucial to discuss the genre of diary novels, as well as single panel cartoons (see Chapter 4), and their narratological implications in order to provide a solid basis for the subsequent analysis.

3.1 A Fiction of Factuality and Intimacy

A key characteristic of diary fiction is its imitation of keeping a real-life diary, which typically presupposes that the entries are intended only for the diarist's eyes. Kellner describes this tension by proposing a threefold notion of fictionality in diary novels: a fiction of intimacy, a fiction of the editor and a fiction of factuality (Kellner 254).²²

First, the diary novel imitates a real diary, which presupposes an intimacy between the diarist and his:her diary, opening a space where he:she can express his:her thoughts freely with no audience. When the fictional diary copies the real one's functions, a sense of fictional intimacy is created. This fiction can, however, be mocked by an ostentatious meta-reflection of the supposed intimacy in the form of literary augmentation, which reveals the fiction (Kellner 255). Further "the fiction of privacy collapses the moment [the diarist] explains his[:her] existential circumstances to

²² In the German original, Kellner uses the terms "Intimitätsfiktion," "Herausgeberfiktion," and "Faktualitätsfiktion" (254).

him[:her]self in the manner of an autobiographer addressing future readers” (Cohn 208), since an explanation of, for example, who other characters are is obsolete if there is no addressee. After all, it is not the diarist who records notes for him:herself exclusively, but an author writing a novel from a fictional character’s perspective, meant to be sold to an abundance of readers.

Second, there is the issue of how a supposedly real-life diary would become published and thus available to a wide audience. Sometimes the real-life author acts as the editor, who publishes the notes by the diarist, which can be indicated in a preface, epilogue, or other editorial notes, thus, creating the fiction of an editor (Kellner 257). For the novels analyzed in this thesis, a fictionalization of an editor is not relevant, however, and will thus not be elaborated on further.²³

The fiction of factuality also constitutes itself through the adherence to private, non-commercial diaries as the real-life model for fictional diaries, copying the narrative processes. Given the fact that these real diaries comprise factual events that happened to factual individuals, fictional diaries adopt this claim to factuality. However, Kellner notes instances, where the fiction is tangible. This is the case when the novel’s composition becomes apparent, for example when the transition from one entry to the next seems too seamless for a character without authorial aspirations or skills.²⁴ Further, an exaggerated ability to remember events and impressions from the past can be another indication of fictionality, when for example vivid visual impressions and the emotions they evoke are described in graphic detail even years after the event. Further, incidences of narrative metalepsis imply fictionality. These occur when the narrative instances, the narrating and experiencing I, become apparently intertwined and there is no clear distinction between author, narrator (=diarist), and fictional character (258–60). Another risk for abandoning the fiction of factuality is if the fictional diarist is not an author or writer and overt literary means implemented in the diary novel by its real-life author disillusion the reader. However, Field notes that

²³ For a more detailed discussion of this topic see Kellner or Chatman.

²⁴ Kellner gives the example of a diarist, who starts his entries by reflecting on his writing process and then seamlessly takes up the narration where it left off in the last entry (259). This would require a lot of literary skill that cannot be assumed for all fictional diarists, especially in the context of children’s or YA literature.

although a fictional diary novel may contain obvious literary devices, this does not automatically reduce its credibility as being the non-fictional thoughts of a non-fictional character (102).

Kellner notes that although indicators of fictionality are predominant in diary novels, there are also instances where elements of factuality are employed intentionally, for example when the fictional world overlaps with the factual world of the author, like when fictional places strongly resemble places from the real-life experience of the author or even are specific factual locations. She concludes that not specific individual elements determine whether the illusory factuality is validated or disturbed, but that this is shaped by the context (259–60).

3.2 Narrative Stance in Diary Novels

Narrative stance comprises key narratological categories such as time, voice, focalization, and subjectivity, which were coined by the French literary theorist Gérard Genette. This thesis examines how these categories manifest in a multimodal context, and the following sections will discuss them in relation to diary novels and, subsequently, cartoons, and later also the multimodal genre of cartoon-diaries. While this thesis acknowledges that these categories stem from literary studies and therefore primarily address verbal narratives, they have also been fruitfully applied in comics research to explore their multimodal implications. This thesis follows this approach in order to enable a systematic analysis of the hybrid genre of cartoon-diaries, beginning with diary novels.

As far as time is concerned, the diary novel foregrounds discourse time (i.e. the time of writing) over story time (i.e. the time written about) because the act of writing is often salient (Kellner 238; Martens 4–5). The prominence of the written word as the medium for telling an individual's story strongly shapes the genre of fictional diaries, as there is a profound difference between dictating a day's events into a dictaphone and expressing thoughts in writing, which implies an even higher amount of privacy (Bauer 6).

The reports in the diary are characterized by an immediacy of events and stances thereto (Martens 4–5). This immediacy only accounts for the character, not the story, as there is always a delay between the events as they occur and their transcription by the diarist. This gap, according to Chatman, is what separates fictional diaries (and epistolary narratives) from truly contemporaneous forms such as the interior monologue (171). Further, it allows the diarist to express his:her stance towards the events described and reflect on his:her, as well as other's, actions, whereas a simultaneous record would be limited to the emotions of the situation. Thus, diary novels are considered to be a plausible method of letting the readers experience a character's, i.e. the protagonist's and implied author's, thoughts similar to other narratological techniques such as free indirect speech or stream-of-consciousness (Martens 189). When the act of writing itself is explicated, the story and the discourse time become intertwined. The focus on the discourse time enables characterizing the narrator, i.e. the diarist, through his:her writing process (Kellner 238), for example, how many entries are written per day, the (ir-)regularity of the entries in length, or the intermittency of breaks in between entries (Bauer 9–10). The discourse time further becomes a level of summarizing, making the diarist a chronicler, who structures and recounts the (immediate) past. Since, generally, the discourse time is prioritized over the story time, there is often more involvement with the diarist as a writer rather than his:her role as protagonist in the story.

Further, how the diarist uses analepsis and prolepsis to talk about events can be crucial for his:her characterization (Kellner 239). The diarist, according to Chatman, has the same knowledge about the future as the reader, as he:she recounts events from the immediate past, him:herself being in the middle of the story, unaware of which events are going to be of relevance, approaching an unknown future together with the reader (171). Genette refers to this as an “interpolated” narration, which takes place “between the moments of the action” (217). The pace in which the diarist narrates can also vary and serve as a means of characterization. A slow convergence of the story and discourse time might indicate the diarist's need for reflection of the recounted events, whereas a slow pace in analepses might try to legitimize present events, or escape back into or relive the past (Kellner 240).

Mode as a narratological category, rather than a semiotic channel, refers to a story's immediacy, focalization and perspectivation. Diary novels create distance not only through the aforementioned time gap between the narration and the events themselves, but also by employing narrative mode (where a narrator is palpable, i.e. telling). By means of narrative mode, the diarist can distance him:herself from the events and the time they occurred. However, dramatic mode (i.e. showing) can also be used, which creates a higher degree of immediacy (Kellner 243–44). Chatman argues that, in diary novels, the story does not necessarily need to be progressed solely through narrative mode, but that the diarist can also convey the immediacy of present affairs from his:her perspective, thereby creating a strong impression of a specific moment in the story (170).

Typically, there is an internal focalization in diary narration, meaning that the narrator and the protagonist have the same amount of knowledge (Kellner 245). However, it could be argued that the narrating I is more reflected and has, apart from gaining knowledge of events in between the entries, also developed a more (temporally) distant stance to the story. In contrast, the experiencing I is a part of the story and functions as an actant within it, possibly having different emotions due to his:her immediacy to the events. Genette notes that in fictional autobiographical texts the experiencing I and the narrating I, or, in his words, the hero and the narrator, can never meet and eventually conflate. He explains it as follows:

“So it is necessary that the narrative be interrupted before the hero overtakes the narrator; it is inconceivable for them both together to write: The End. The narrator's last sentence is when – is *that* – the hero finally reaches his[:her] first. The interval between the end of the story and the moment of the narrating is therefore the time it takes the hero to write this book, which is and is not the book the narrator, in his[:her] turn, reveals to us in a moment brief as a flash of lightning.” (227, original emphasis)

Further, since the narrator of the diary novel is also the story's protagonist, the narrative voice is autodiegetic and thus clearly individualized. Apart from occasional moments in which the act of writing is explicitly reflected (i.e. the present), the narration typically recounts events retrospectively. In most cases, the narrative voice remains that of the diarist; however, when letters, notes, or other inserted documents are included, additional intradiegetic voices may emerge within the diary (Kellner 246, 249).

Given the prominence of the diarist, the question of who is talking, i.e. narrative voice, can be easily and exhaustively answered, the issue of who is addressed proves to be more complex. Therefore, Bauer refers to the paradox of the addressee. Although narrative fiction is based on the principle of communication, which entails some form of a receiving end, i.e. an audience, a non-fictional diary is usually not meant to be read by anybody apart from the diarist him:herself. This does not mean, however, that an addressee is missing, since most diary entries are directed towards somebody or something, which is not necessarily a reader (Bauer 60). Forms of an addressee could include an imaginary “you,” an idea inside the diarist’s mind. Martens notes that above all, a diarist writes for him:herself, being the sender and receiver in one person. Field describes the diarist in a double function: first, as the author, and second as a different person who simultaneously functions as the audience. Sometimes it is the diary, which is the addressee itself, becoming a personified imagination that exists in the diarist’s mind, as is often expressed in the archetypal phrase “Dear Diary” (Bauer 70).

3.3 Diarists – Narrators and Protagonists in Diary Novels

Apart from the defining narratological frame such as first-person narration and analepsis, as well as topical conventionalities such as loneliness or “quest for self or affirmation of self” (Prince 497), it is above all the theme of writing a diary that sets fictional diaries apart from other novels. This entails a reasoning of why the diarist starts recording events in a journal, as well as, to some extent, references to the physicality of the diary or its dialectical relationship with the diarist. Thus, the diarist has a constitutive role for the genre. Often the diarist writes entries when he:she experiences some sort of crisis, making the diary a means of reflection and self-defining (Kellner 223, 228–29). This suggests that it is less a given formal structure, such as a division into sections with specific dating, which defines diary novels as a genre, but more its subjectivity and meta-narrativity between the poles of fact and fiction (cf. Prince; Kellner), which is channeled through a diarist.

Given the fact that diarists adopt an authorial position towards their own story, analyzing their narrative stance needs to take into account how they portray themselves

as well as others. In linguistically reflective diary novels, the diarist as a subject is tangled up with his:her language and primarily characterizes him:herself indirectly through his:her way of writing. The diarist, thus, is both a writing and written subject and can ultimately decide whether and how to depict him:herself and other characters (Kellner 252). Regarding the fiction of factuality, usually in diary novels, information about other characters, such as their relationship to the diarist, need to be gained from in between the lines, if there is no explicit addressee, for whom such a knowledge cannot be assumed (cf. Field 94).²⁵

The motive why the diarist decided to collect his:her thoughts in a diary certainly is crucial for a characterization. Bauer notes that one main function of keeping a journal is therapeutic in nature, where emotions can be verbalized and reflected within a safe setting. One character trait that, according to Field, most diarists have in common, is a feeling of isolation and loneliness, which compels them to this form of “silent self-communication” (145). For such a character, the diary is a part of their life that is defined by intimacy. Apart from loneliness, Bauer also notes low self-esteem as a possible reason for writing a diary as a form of self-therapy, including the diarist’s search for a way of analyzing and observing his:her own behavior within a space where he:she can focus on his:her own (alleged) shortcomings. To manage this form of self-doubt, the therapeutic function constitutes itself in the verbal expression of one’s own thoughts and feelings and although the writing itself cannot solve any of the diarist’s real-life problems, it helps process events and emotions and might provide some sort of valve for venting. As soon as the situation of the protagonists improves in his:her fictional life outside the diary, he:she might tend to write entries less frequently, as he:she is more occupied with enjoying his:her life (Bauer 90–97).

Apart from these thematic considerations, an analysis of a diarist must also address his:her narratologically split role as both narrator and protagonist, as discussed above. If the two temporally distinct personas of the narrating and experiencing I do not differ, the diarist might still feel in the moment and stand by what he:she said and how he:she acted in a certain situation. Alternatively, the narrating I could critically

²⁵ According to Field, unreliability of the first-person narrator needs to be taken into consideration if there is the explicit intention of publishing the diary.

reflect on his:her past actions and distance him:herself from his:her actions, which would be a key point in his:her personal development. The experiencing I is likely to feel emotions more intensely, whereas the narrating I is able to take an outside stance towards the events and might be more rational and reflected. Whether and how this difference can be made apparent remains an open question, as no direct insight into the experiencing diarist can be gained.

While the narratological foundation of cartoon-diaries is initially shaped by the conventions of diary fiction, as will be discussed later, the incorporation of cartoons significantly modifies these structures. The following chapter therefore turns to narratological considerations specific to cartoons in order to account for these shifts.

4 Cartoons – Narrative Implications of a Multimodal Genre

Cartoons in themselves are multimodal artifacts, as they comprise both a typical form of iconic imagery as well as textual elements in, for example, speech balloons or captions. While cartoons, usually only comprising single or minimally sequential panels, are distinguished from comics mainly by their length, they share with comics their reliance on iconic imagery and the ways in which they unfold their narrative potential. Both genres mobilize the interplay of image and text to construct meaning, evoke temporality, and shape perspective. For this reason, seminal works of comic studies will inform this chapter.

To start off, it is necessary to consider how multimodality is configured in cartoons. Regarding the interplay of verbal and visual modes, Scott McCloud identifies seven major word-image relationships in comics. Given the similarity to comics, cartoons follow the same principles. While some cartoons may rely predominantly on a single mode, whenever both words and images are present, they can interact in different ways. First, pictures can illustrate a text, without vitally contributing to the meaning, which McCloud refers to as a word specific combination. Conversely, in picture specific combinations, words are largely absent or function only as a “soundtrack to

a visually told sequence” (153). In duo-specific combinations, both modes essentially communicate the same content. Further, an additive relationship occurs when one mode amplifies or elaborates the other, whereas in a parallel combination, words and images follow separate, seemingly unrelated paths. A montage integrates verbal elements directly into the visual composition itself. Finally, what McCloud sees as the most common form is the interdependent combination, in which words and images jointly construct meaning that neither mode could fully convey on its own (153–55).

Possible configurations of an interaction between verbal and visual modes in comics have also been described by Packard et al., whose model can equally be applied to cartoons. They describe the connections that readers establish between elements in a multimodal artefact as semiotic or intermodal relations (60) and identify five principal types. The first, extension, refers to the specification of one element through another. Second, a part relation denotes a part-whole relation between elements. Third, contrast is characterized by structural similarity but semantic dissimilarity. By comparison, parallel describes cases in which elements display both structural and semantic similarity. Finally, narration refers to a temporal-spatial succession of elements revolving around a shared topic (61–62). These relations may occur individually or in combination, resulting in more complex multimodal structures. Packard et al. therefore raise the question of whether such layered structures are potentially more interesting and effective than a simple narrative structure (69).²⁶ Whether one adopts McCloud’s typology or that of Packard et al., it becomes clear that the interaction of modes in multimodal artefacts is highly diverse and must be examined individually for specific cases, which will be undertaken in Chapter 6.

²⁶ They elaborate on these relations based on McCloud. They use his example of wavy lines signifying smoke coming from a pipe, which they define as an extension of the pipe as it signals that it has been lit. Later, McCloud uses the same lines above a garbage can, where instead of smoke, they now signal an unpleasant smell. Given the structural similarity, the readers would here infer a parallel between the two uses of the lines, but given the semantic dissimilarity would infer a contrast in the way the lines are contrasted (Packard et al. 61).

4.1 Iconicity of Cartoony Imagery

In order to establish a theoretical basis for the discussion of iconicity, it is useful to return briefly to semiotics. From this perspective, Charles Sanders Peirce's concept of the iconic sign is particularly relevant. He defined an icon as a sign that bears sufficient visual or structural resemblance to its object for the interpreter to recognize the connection (cf. Bateman et al., *Multimodality* 59). This understanding of iconicity has since become a central category in comics studies. For instance, the influential comics theorist Scott McCloud describes icons and iconic imagery as the fundamental building blocks of comics, arguing that they constitute the medium's basic vocabulary. Icons, for McCloud, constitute the abstractions of real-life objects. The degree of abstraction varies along a spectrum from highly realistic drawings approaching photographic representation, to simplified images, as well as, at the most abstract end of the spectrum, ultimately words. In this sense, McCloud considers both pictorial and verbal elements in comics to be icons, ranging from a detailed drawing of a face to a minimal smiley or even the written word "face," all of which constitute forms of abstraction (McCloud 27–29). The increased abstraction of cartooning does not necessarily eliminate details, but highlights specific elements, revealing and subsequently augmenting the essential ideas, inducing "a form of amplification through simplification" (McCloud 30). Also, Eisner refers to the process of cartooning as "exaggeration and simplification" and sees the cartoon as "a form of impressionism" (163). By eliminating nonessential details, the focus can be shifted towards the essence of meaning. Cartoons focus the attention of the reader on an idea rather than a subject, thus inscribing a form of universality rather than specificity. While highly concrete depictions can foreground an object's physical appearance and highlight specific details about it, an iconic cartoony style conveys a more general idea of the same object (McCloud 41).

Building on the idea that cartooning emphasizes the essential through abstraction, McCloud also explores the reception process involved in reading comics and cartoons, which he refers to as closure. Closure describes the act of "observing the parts but perceiving the whole" (63). When a few simplified, cartoony elements are arranged into a recognizable image, the reader naturally completes the picture, even

when visual information is missing or only implied. This act of mental completion occurs within some form of visual boundary, whether this be a literal frame, the edge of the page, or even an unmarked space. What lies outside of the (in)visible line is open to the imagination of the reader, although this is not defined by arbitrariness, but follows learned processes of closure. For example, when the reader is presented with an illustration showing a person from the waist up, he: she still assumes that legs are attached to that person, despite the lack of explicit portrayal thereof (McCloud 61). The mind processes the lines depicted on a piece of paper, recognizes the abstraction and loads it with real-life experiences to receive two dots with a line underneath (i.e. :)) as a human face. Through this process, according to McCloud, with a higher level of iconicity the reader involvement increases as well.

The process of closure also relates to deciphering visual metaphors, which are another key aspect of iconic depictions. These simplified images often stand in for more complex concepts, requiring the reader to engage in interpretive work. McCloud offers an example of this, illustrating how a few wavy lines can be interpreted as smoke rising from a lit pipe. For this interpretation to take place, the reader must rely on prior knowledge, i.e. that when a pipe is lit, it emanates smoke, and recognize that the lines, albeit abstract, represent an element that is not directly pictured (128). Interpreting a visual metaphor thus relies on learned meaning-making process, which are also highly dependent on the context. The same wavy lines, when placed above a garbage can, may be understood not as smoke, but as an unpleasant smell. This flexibility in meaning illustrates the polysemy of visual signs in cartooning and highlights how closure, guided by convention and context, enables readers to make sense of abstracted visual cues (cf. Packard et al. 52–58).

A rather prominent visual metaphor essential in comics are speech balloons. These represent sound, or more precisely speech, but the words they contain are irrelevant for the visual metaphor. More so, the lettering of the words and the shapes of the balloons themselves carry meaning and convey qualities such as volume, tone, or intensity, materializing something as ethereal as speech. Variations in form further differentiate types of utterances, such as cloud-like balloons typically indicate thoughts, whereas zigzag outlines pointing toward a device imply electronically transmitted

speech (Eisner 24–25). In this way, speech balloons function as conventionalized visual metaphors, whose meaning depends on culturally learned interpretive practices, thereby contributing to the recognizable visual vocabulary of cartoons.

Beyond abstraction, closure, and the use of visual metaphors, cartoon-specific iconicity entails a range of further aspects that cannot be addressed here without exceeding the scope of this thesis. For the present purposes, however, these features represent the most relevant characteristics of cartoons as a pictorial mode. Building on this foundation, the next question concerns how narrativity is induced within this multimodal genre, which will be discussed in the following section.

4.2 Narrativity of Cartoons

In general, narrativity does not rely on the presence of verbal input alone, but is possible in any medium (Ryan). Baetens states that humans tend to read even “non-figurative material in a narrative manner” (101), which makes even more abstract artefacts narratively interpretable through human’s “ability to narrativize” (Wolf 271).²⁷ Given the presence of, what Wolf calls, “core narremes” (265), such as a storyworld that is situated in time and space and is inhabited by characters, narrativity is induced also in non-verbal input. Since cartoons have no problem fulfilling these narremes, their potential to generate narrativity is beyond dispute and accordingly, the narratological categories time, mode, voice, and focalization will be outlined below.

Firstly, cartoons can represent time and, in doing so, generate narrativity. Similar to comics, this temporal dimension is often conveyed through the sequential arrangement of images, allowing events to unfold across space in a way that suggests progression. Since ‘reading’ an image also requires time, i.e. discourse time, or, in Wolf’s terminology for comics, “reception time” (265),²⁸ a form of temporality is inevitably involved, however, this should not automatically be equated with narrative temporality. Indicators for narrative temporality are visual cues, gestures and overall

²⁷ Wolf argues that artifacts such instrumental music, sculptures, poetry or dance can also fulfil a narrative communicative function (276).

²⁸ Story time would then be termed “represented time” (Wolf 265).

composition, which can be present even in single-panel cartoons and guide the viewer's understanding of temporal and causal relationships.²⁹ Even recurring single-panel cartoons can have narrativity, if they connect over time (cf. Abdel-Raheem). Additionally, time can be expressed linguistically, through narrative captions or speech balloons, which are themselves inherently temporal: Speech, or more precisely sound, in itself is not instantaneous, but exists only in time, rendering only silent panels truly instantaneous (McCloud 95). Irrespective of whether the panel is silent or not, it can be comprised of several moments in itself, where action and reaction are depicted simultaneously, but actually show a temporal succession. This becomes especially evident when (verbal) interactions are depicted within the same panel, where utterances are present simultaneously, but logically lag behind each other in a question-response format, as well as accompanying bodily reactions such as facial expressions or gestures (McCloud 95). The co-presence of sound and reactions is also evident in the case of onomatopoetic sound effects (McCloud 116).

In addition to sound, the representation of motion constitutes a second means by which single-panel cartoons can express temporal progression. Over time, artists have developed conventions such as motion lines to visualize movement within a static image. These lines may vary in complexity and stylistic realization across cultural contexts, yet they function as a widely recognized visual metaphor for movement, enabling viewers to infer dynamic action from a still depiction (McCloud 109–15). Additionally, Packard et al. show how time can be progressed or stopped through other visual cues, such as the gaze of characters.³⁰

This interplay of visual and verbal temporality allows cartoons to convey stories, character actions, and events, establishing their capacity for narrativity even within seemingly static or minimal formats. Thus, even non- or not-exclusively verbal artefacts,

²⁹ For an in-depth discussion of how temporal progression is expressed in comics see McCloud (94–117) and Eisner (23–38).

³⁰ They give an example of a comic, where the protagonist is looking to the right, thus progressing the story, while side characters gaze to the left. Packard et al. call this lateral gazing. In contrast, frontal gazing can mean that time in the storyworld is paused and a character addresses the readers directly (32).

such as cartoons, fulfill a narrative communicative function (cf. Page; Bateman et al., *Multimodality*; Kukkonen, “Transmedial Narratology”; Doloughan; Wolf; Thon).

Secondly, with regard to narrative mode, comics and cartoons differ significantly from purely verbal narration. While both narrative (i.e. telling) and dramatic (i.e. showing) modes can be realized in comics, the visual dimension foregrounds what has traditionally been described as showing. In film theory, showing is conceptualized as ‘monstration,’ which anchors actions in the present. Even when past events are represented, they are depicted as a then-present moment of action. A narrative past or future then only emerges through the editing of a movie rather than a single image itself, resulting in a moment in the present, but a story in the past (Gaudreault 32). As opposed to movies, monstration in cartoons is more tangible, as the graphic style constantly foregrounds its own constructedness, making their reception less immediate than in a cinematic or photographic depiction, anchoring the story even more in the past (Mikkonen 86–87). Groensteen (*Narration* 82) further emphasizes that comics tend to orient narration toward the past, as previously viewed panels remain visible and accumulate meaning. Unlike film, where images replace one another, comics preserve earlier moments, giving readers greater control over reception and reinforcing the retrospective character of the narrative, since “unlike the past, the present cannot be cumulative” (*Narration* 87). Although an isolated image can only show, the arrangement of multiple images enables storytelling.

Thirdly, the concept of narrative voice, i.e. the presence of a voice distinct from the author (Genette 213), is complicated in non-exclusively verbal media, where such a presence is not necessarily required (cf. Packard et al.; Thon; Mikkonen). Particularly in pictorial narratives, Thon suggests the concept of a “non-represented verbal-pictorial narrating instance” (105). Such a graphic enunciation approach allows the readers to accept that the representation of the storyworld does not have to be made by a narrating instance, thus making a narrator an optional strategy used in (visual) narratives. Accordingly, a narratological analysis of non-verbal or multimodal artefacts must first ask, whether a narrating instance can be identified and, if so, how it is positioned within the narratorial hierarchy, how it relates to the storyworld, and whether it

tells the story, verbally, written, or mentally (Packard et al. 83). In comics, narration may occur verbally, through captions, word balloons, or pictorially, through the graphic presentation of events without explicit verbal mediation. Groensteen differentiates between various functions in this process, and posits, for example, the narrator as the one that organizes and structures the narrative information, while the monstrator visualizes characters and actions. If there is an autodiegetic narrator, he:she can either choose to use first person view and thus stay in the background, or to show themselves and make themselves tangible. When represented, this narrator becomes part of the graphic world and is, like all other characters, subject to the same processes of visual construction by the monstrator (*Narration* 90–99).

Regarding diegesis, Thon proposes that in visual media, such as comics, the distinction between extradiegetic and intradiegetic narration depends on whether the narrating instance is represented within the storyworld. Narration is intradiegetic when the narrator appears as a character and the act of narrating is depicted as part of the primary diegetic level. Extradiegetic narration, by contrast, operates from outside the storyworld and is not represented as part of it, typically remaining invisible or unacknowledged by the characters (107).³¹

Lastly, the expression of subjectivity, or focalization, is central to narratives and for pictorial narratives specifically, Thon argues for a broad conception of a “representation of (some form of) subjectivity” (110). In comics and cartoons, subjectivity does not have to be expressed via captions or thought balloons, but it can also be conveyed pictorially, which relies heavily on the use of visual metaphors that externalize a character’s inner states (Packard et al. 89, 121–31). However, the visualization of subjectivity has to be cued, as showing a subjective perspective is not the norm in pictorial narratives, according to Mikkonen. Kukkonen similarly argues that images “are usually perceived for what they show, not for the perspective from which they show it” (*Storytelling* 151–53). This makes the marking of focalization in visual media particularly significant.

³¹ For a discussion on different configurations of narration levels see Groensteen (*Narration*), Packard et al., or Thon.

In multimodal artefacts, focalization becomes structurally complex, as visual theory distinguishes between focalization through a character and through a narrator, linking focalization to optical perspectivation (Horstkotte and Pedri 350). Concerning this conceptualization, Thon distinguishes between objective, intersubjective, and subjective representation. While intersubjective representation constitutes the default mode, objective representation presents the storyworld without anchoring it in a perceiving character, and subjective representation depicts the world as perceived or imagined by a character. As far as the latter, i.e. subjective representations, are concerned, these may range from mind-representations, point-of-view sequences with varying degrees of salient perception, such as hallucinations, or representations of entirely internal worlds, such as dreams (Thon 109–14). If there are shifts between representations of subjectivity, readers have to decipher these by recognizing a range of subconsciously established subjectivity markers (Packard et al. 90–94). Ultimately, questions of focalization in multimodal narratives raise the issue of intersubjective validity, depending on how far a depicted perception can be regarded as reliable within the reality of the storyworld.

4.3 Multimodal Characters

Typically, narrators are characters from within or without the storyworld. In pictorial narratives, however, characters are typically presented as part of that storyworld. Therefore, to conceptualize a graphic narrator, it is necessary to examine how character representation works in the cartoon medium.

As mentioned above, the presence of characters contributes to a (multimodal) text's narrative potential. In pictorial narratives, especially through dramatic mode, characters often serve as central agents of storytelling. It is particularly the “continuing presence of a character or a group of characters in a sequence of images” (Mikkonen 90), which is a constitutive feature of narrativity in comics/cartoons. While characters are not strictly required for narrativity, recurring figures provide cohesion, guide focalization, and offer readers a stable point of orientation. Especially changes in a recurring character's looks or behavior indicate a temporal progression and move the story along (Mikkonen 96). Through facial expressions, gestures, positioning within the

frame, and exchanges of gazes, characters create spatial, temporal, and experiential connectivity (Mikkonen 83, 106). Characters are thus not only elements of the story-world but also structural devices within the narrative composition. In general, to qualify as a character, a figure (in human, animalistic, or other shape) must display agency, i.e. the capacity to initiate and take responsibility for actions (Aldama 319). Agency and intentionality therefore distinguish an inanimate object like a rock from a character in the shape of a rock, for example.

The inherent iconicity of the cartoon medium further shapes the representation of its characters and potential narrators. Cartoon figures are typically shown in an iconic style, which means that they are shaped through varying degrees of abstraction. In contrast to a realistic depiction, where a character's individual physical details are highlighted, an iconic shape highlights the conceptual and more general aspects of a character, by focusing on the idea behind it. This is what McCloud calls the difference between portraying "the world without" and "the world within" (41). A highly abstract face, for example ☺, is so universal it could describe nearly all human faces, thus inviting broad identification, whereas a realistic depiction, for example a photograph, foregrounds individuality, and represents only one person. The level of iconicity therefore influences both reader identification and interpretation, without implying reduced character complexity.

Iconicity further influences the depictions of emotions and mental states, which are likewise conveyed through a repertoire of visual conventions. Some depictions derive from observable phenomena, like sweat beads signaling nervousness. These visible drops have been abstracted and conventionalized as metaphors over time and have gained a symbolic quality, which McCloud terms a "drift from visible to invisible" (130). Cartoons and their iconic characters and narrators thus rely on a learned visual vocabulary of metaphors, gestures, and facial expressions to externalize inner states.

Apart from this form of visual depiction, which also relies heavily on the use of visual metaphors, multimodal characterization in cartoons also emerges through dialogue, captions, and layout (Mikkonen 176). Obviously, external traits, such as a character's

appearance are rendered visually³², while the choice of style that is used can trigger connotations, such as triangular shapes might denote sinisterness (Aldama 323). Feelings and mental states can be expressed both visually, through facial expressions, gestures, posture, or visual metaphors, or verbally through thought balloons (Kukkonen, *Storytelling* 151). In this way, visually based media can convey no less complex inner states than those relying on strictly verbal modes. While in verbal-based novels the meaning of utterances is contextualized by a description of undertone or posture, cartoons can portray this context visually. The words “I’m sorry” can therefore be uttered by a person on his:her knees or by someone with an arrogant facial expression that question the sincerity of the apology (Eisner 106). These postures, gestures, and facial expressions are part of a learned dictionary of idiomatic expressions humans use. How the anatomy of the human body is shaped therefore has highly expressive potential. The human face, in particular, functions as a focal point of reader engagement, often serving as a “window to the mind” (Eisner 114).³³ Unlike prose, comics do not require verbal tags such as “she screamed,” however, a visual depiction of a character yelling at another can relay details, such as emotions, of the situation visually, which is perceived instantaneously (Aldama 322).

In sum, cartoon characters combine agency, visual stylization, and multimodal expression to generate narrativity. Their representation draws on a partly conventionalized visual vocabulary that enables them not only to inhabit the storyworld, but also to structure and progress it.

³² This spectrum ranges from more detailed representations to highly iconic, abstract, or even metaphorical depictions. As Aldama points out, visual characterization often signals traits such as gender or race, but there are also deliberate deviations from this practice, for instance when characters are drawn in a visually uniform manner and are marked as “Asian” primarily through their names (Aldama 320–21). A further example of metaphorical characterization can be found in Art Spiegelman’s seminal graphic narrative *Maus*, in which Nazis are represented as cats and Jews as mice.

³³ Expressing complex thoughts in cartoons and comics might only be limited by the space that is at the artist’s disposal. Eisner, for example, rendered Hamlet’s notorious monologue in a comic format, which takes significantly more space than the mere words in a play, but is therefore enriched by the character’s posture, gesture and facial expressions (Eisner 115–24).

5 Narrativity of Multimodality – Cartoon-Diaries

Much like other multimodal genres such as comic books, graphic novels, or picture books, cartoon diaries combine verbal and visual modes and can thus be described as hybrid or “fusion” texts (cf. Evans). Within this format, however, verbal narration generally remains dominant, while the cartoons are embedded within stretches of running text rather than juxtaposed to each other. Still, the iconic images contribute significantly to the progression of the narrative by visualizing events in ways comparable to comics. Since McCloud classifies even single-panel cartoons as a form of “comic art” (20), this chapter draws on insights from comic studies, especially with regard to narrativity. The following section therefore outlines the genre of the cartoon-diary and discusses the implications of its multimodal form on its narratology, thereby bridging the theoretical foundations established so far with the close analysis in Chapter 6.

5.1 Genre as Framework: The Diary Novel in a Multimodal Context

Genres shape expectations, establish conventions, and provide a framework for comparative analysis. For this reason, they constitute an essential category in multimodality studies as well (cf. Bateman et al., *Multimodality*; Kress). As Bateman et al. argue, a text cannot be separated from its genre, since genre fundamentally informs how meaning is produced and interpreted (*Multimodality* 129–30). This is particularly relevant in multimodal contexts, where genres are often hybrid and combine different ways of meaning-making (Kress 106–07, 112). With that in mind, a description of the genre of the cartoon-diary will be outlined below.

To account for the genre’s multimodality, the canvas as the site of modal interplay must first be outlined. In this case, the canvas is the printed novel: a two-dimensional, permanent, (usually) black-and-white medium consisting of pages to be turned. It draws on the real-life counterpart of the personal diary, appropriating and fictionalizing its conventions, including its material qualities. Accordingly, the absence of a handwritten font, which is usually associated with actual diaries, might be considered a deviation from this fictionalized authenticity. With the canvas thus

established, the following sections turn to the modes themselves and their respective contributions.

Written language in cartoon diaries fulfills several functions. First, it advances the story by fulfilling a narrative/telling function. Second, it largely determines the overall genre framework of the text as a diary novel, for example through elements such as dated entries or phrases like “Dear Diary.” Third, as a visual mode, written language also affords semiotic resources such as typography, where the use of a handwritten-style font, for example, can reinforce the fiction of factuality. Fourth, it can establish “multimodal cohesive ties” (Stöckl 55) by referring to the iconic images or referencing the act of drawing itself. Finally, written language also appears within the inherently multimodal cartoons, for example in speech balloons, captions, or dialogues.

Iconic imagery contributes additional functions. First, the simplified, iconic style can enhance reader involvement, as abstract depictions allow for a projection of the self and facilitate identification with the characters (cf. McCloud). Second, images show the characters, settings, and objects of the storyworld in a visible form, thereby fulfilling an important showing function. This showing function of the iconic imagery can even strengthen the assumed non-fictionality of the diary by materializing characters and attributing defined physical features, thereby linking fictional characters and objects to an assumed real storyworld. Third, by visually presenting the storyworld, the cartoons can also create a distance between the narrating and the experiencing diarist, who is externalized from the usually introspective visual representation typical of diary novels. This separation has important narratological implications for the genre and will be discussed in the following sections.

Generally, the ratio between verbal and pictorial elements in cartoon diaries can vary, ranging from minimal or non-linear text to sparse cartoons that either establish their own narrative or merely serve illustrative purposes. To account for these variations, a prototypical approach is adopted here (cf. Packard et al.; Wolf; Bateman et al., *Multimodality*). In the genre of cartoon-diaries, and specifically in the two works analyzed in this thesis, verbal narration generally dominates, shaping the narrativity primarily through the “diary” aspect of the genre, thus focusing on the stem word of the composite term cartoon-diary. Since *ATD* and *DWK* are highly influential and

widely popular titles that have defined the genre and inspired numerous imitations, particularly in children's and YA literature, their configuration proves a suitable basis for a prototypical approach.

Thus, diary narration remains the defining feature of the genre, while the inclusion of cartoons introduces a multimodal dimension that gives rise to the specific sub-genre of the cartoon diary. This multimodal configuration has two major implications: a splitting of the narrating instance and a shift in the temporal focus compared to conventional diary narration. These two aspects will be examined in greater detail in the following subsections.

5.2 Temporal Structures in Multimodal Diary Narratives

In traditional diary narration, discourse time typically is foregrounded over story time. In cartoon-diaries, however, multimodality produces an interesting shift, in that by including imagery, i.e. showing mode, this focus on discourse time is softened. While the diary structure continues to organize the narrative verbally, the integration of images introduces an additional level of temporal representation. Because written language unfolds sequentially in time whereas images are spatially organized, their interaction leads to a partial spatialization of narrative time, as well as a temporalization of spatial images. These dual effects on temporality, i.e. both the narratological and organizational implications, will be explored in the following section.

First, the softening of discourse time will be discussed. In diary narration, usually, the discourse time, i.e. the time of writing, is foregrounded over the story time, i.e. the time written about. Typically, the verbal text in diary novels conveys events from the past. In multimodal contexts, images shift the focus more toward the present, as, by their nature, they can only show a present moment. When these two temporal planes interact in multimodal diary narratives, they shift the reader's focus from the past-oriented discourse time to the more immediately perceivable present of the story time.

Cartoons in these texts can function in three ways. First, they may depict non-narrative objects, or people related to past events. The verbal narrative can situate the image in the past, present, or future, yet the snapshot in itself is not very temporal, making the conflation of past and present insignificant. Second, images may be subjective, portraying imagined futures, dreams, or surreal scenarios. Here, the visual content introduces a hypothetical or non-existent present, and the multimodal context determines whether what is shown should be interpreted as real or imagined (cf. Bateman et al., *Multimodality* 120–21). Third, cartoons can depict actual scenes from the diarist's life that are reflected upon in the narrative. These images show a moment of the present within the scene, while the verbal narrative situates it in the past, often as an analepsis, which is typical for diary narration. This juxtaposition of past and present causes a stop of the discourse time, sucking the reader into the story time, emphasizing the immediacy of its state as “having been the present” within a predominantly past-focused discourse.

It is exactly this interplay of shifting from the discourse time to the story time that makes the diarist's experiences more accessible and vivid for the reader. In multimodal diary novels, cartoons therefore contribute a present-oriented layer through the images to a past-centered verbal narration, by switching the narratological mode from telling to showing, introducing a subtle disruption to the typical temporal structure of diary novels.

Beyond these narrative effects, multimodality in cartoon-diaries also influences the genre's organization, as previously discussed in Chapter 2.2.3 in terms of a spatialization of time and a temporalization of space (cf. Doloughan 34). In this genre, cartoons are interspersed across a verbal narrative, producing a similar effect. As Doloughan observes:

“The assumption of the dominance of time in narrative, in terms of progression, as well as in relation to the experiencing of events, actions and behaviours, is undercut by a disruption, even a suspension, of time. The spatial, which has always been part of narrative in metaphoric and imaginative terms, begins to be given literal extension.” (45)

The verbal narration is organized temporally, one word follows another, one sentence the previous one, and one entry the preceding one. Thus, meaning is created sequentially, both in production and reception. The iconic cartoons, in contrast, are spatially organized, which makes them perceivable as a whole, and they are interpreted through spatial relationships. In cartoon diaries, the insertion of images into the verbal narrative softens its temporal dominance, as the written narrative is spatially interrupted, allowing the reader to navigate the page freely, choosing whether to read the text first or the cartoon. Consequently, the spatial layout of the page shapes the temporal experience of the narrative.

5.3 The Multimodal Diarist

The distinction between discourse time and story time further reinforces the split between the two narrating instances in diary novels: The narrating I occupies discourse time, while the experiencing I acts within the story time.

In autodiegetic diaries, the diarist thus functions simultaneously as both narrator and protagonist. While this separation exists in conventional diaries, it becomes more distinct in multimodal cartoon diaries, where the I-as-narrator can also visualize him:herself. The narrating I structures the narrative using literary devices³⁴, references the act of drawing, and positions him:herself as writer, cartoonist, and storyteller, while the experiencing I is materialized verbally or graphically through past events being recounted. By spanning two modes, the split of the diarist becomes tangible, allowing for a more nuanced evaluation of the diarist's actions and personality. Readers gain insight not only through verbal self-descriptions or textual cues but also through the diarist's self-depiction, which can be simultaneously revealing and unreliable, as will be seen upon analyzing *DWK* in Chapter 6.2.

³⁴ Clearly, the examples, as fictional novels, do employ literary devices and narratological strategies, but they should not be apparent to the reader, as it would not be credible for both protagonists if they were able to command these in a sophisticated way. To uphold the fiction of factuality and intimacy, the literary devices need to be employed to give the reader a pleasant experience, but cannot be apparent.

In cartoon-diaries, the diarist is both narrator, protagonist, as well as monstrator, so both the written and visual modes generally present a subjective perception of the storyworld. However, the degree of subjectivity can vary. For example, if cartoons are produced alongside the verbal reflection of past events, they are focalized through the narrating I rather than the experiencing I, representing narratorial rather than character-bound focalization. As Horstkotte and Pedri note:

“Therefore, graphic memoir may introduce gaps and lags not only between the experientiality of the experiencing-I and its retrospective reconstruction by a narrating-I, which is necessarily tinged by the aspectuality of the latter, but also between the two semiotic tracks” (350)

Further, a temporal gap between the narrating and experiencing I enables a possible development of the diarist, allowing the narrating I to reflect on and re-evaluate past events (cf. Horstkotte and Pedri 338–39, on *Persepolis* by Marjane Satrapi). Groensteen further argues that “drawing oneself necessarily involves distancing oneself” (*Narration* 99), as this self-externalization makes the diarist the protagonist of his:her own story. In cartoons, characters are shown in the present but, from the perspective of the narrating I, represent the past, visually reinforcing this distance. As monstrator, the narrating I controls how these depictions are enunciated graphically. Conversely, if the cartoons are created closer to the narrated events, they are focalized more through the experiencing I. There are also instances where either mode may present material that is not focalized through the diarist, for example a letter from another character included in the diary, which conveys a character-bound perspective that the narrating I can evaluate. This interplay of narratorial and character-bound focalization, across verbal and visual modes, is a defining feature of multimodal diary narratives.

The multimodal reinforcement of the split between the narrating and experiencing I entails further implications for characterization. Since both the writing and the drawing processes are explicitly linked to the diarist, who is transformed through the act of diary-keeping just as much as through the story itself (Kellner 238), the character can be constructed through both verbal and graphic enunciation. Unlike purely verbal narration, where characterization is largely mediated through internal reflection, cartoons can realize both internal and external qualities of its characters

(Mikkonen 197–98). In the hybrid form of the cartoon diary, which traditionally foregrounds inner states such as thoughts, emotions, and attitudes, the visual mode complements and extends this by providing a concrete, albeit iconically stylized, external representation. Since this is a self-representation, it can also give insights into the diarist's character. By revealing how the diarist sees him:herself, indirect characterization takes place. Moreover, the diarist also controls how other characters are portrayed (Kellner 252), and in a multimodal context, this portrayal can unfold across both written language and imagery and also helps characterize the diarist him:herself. The combination of autodiegetic narration and cartoons thus offers insights into both the protagonist's inner life and outward appearance, as well as enables multimodal characterization.

One last aspect of a multimodal diarist is that the hybridity of comics and diary fiction entails an intriguing tension between the internal and external. On the one hand, diary narration is autodiegetic and strongly oriented toward the inner world of the protagonist, i.e. his:her thoughts, emotions, and desires. On the other hand, cartoons introduce an external perspective, particularly through a visualization of the self.³⁵ As a result, the diarist's subjectivity is not expressed solely through language but is also given a visible, external form. This creates a productive ambivalence. While the diary form conveys immediacy and direct access to the diarist's thoughts,³⁶ the visual representations may introduce a certain degree of mediation. At the same time, iconic drawings reinforce rather than diminish immediacy. Unlike highly realistic depictions, which might foreground the individuality and physical specificity of a character, iconic representations remain abstract and thus facilitate reader identification. Although realistic imagery could potentially strengthen the fiction of factuality, cartoon diaries favor abstraction, which enhances accessibility and reader involvement. Varying degrees of iconicity further allow for different levels of

³⁵ Of course comics can also depict an inside view and show what a character is seeing through his:her eyes (e.g. McCloud 40), but this is not the default mode in neither *DWK* nor *ATD*. Additionally, there is a wide array of means to depict inner emotional states externally through visual metaphors or anatomical expressions (cf. Eisner; McCloud).

³⁶ Particularly in children's literature, such as *DWK*, the use of a first-person narrator has proven effective, for this sense of immediacy it creates (cf. Wyile). By contrast, first-person visual narration in children's literature, which in comic studies is referred to as "autographics" (cf. Hughes et al.), is comparatively rare.

engagement, complementing the intimacy created by autodiegetic narration. Ultimately, the internal focus of the diary and the external perspective of the cartoons do not stand in opposition but interact to enrich the construction of meaning.

Having outlined the key concepts of multimodality and narratology relevant to this study, the following chapter now turns to their application. The theoretical framework developed here will serve as the basis for the close analysis of the selected cartoon-diaries.

6 Analysis of Multimodal Narration in Contemporary Diary Novels

Building on the theoretical framework of multimodality and narratology outlined in the previous chapters, the following analysis examines how visual and verbal modes interact in *DWK* and *ATD* to construct narrative stance. As established earlier, narrative stance is understood here as the interplay between narrative voice, focalization, and the narrator's subjective positioning, particularly as these dimensions are shaped through multimodal elements.

Methodologically, the analysis is based on a close reading of both primary texts. Each novel was read multiple times with different analytical foci: first to establish an overall understanding of narrative structure and tone, then to identify recurring multimodal patterns, and finally to conduct a detailed examination of specific image-text relations. The analytical framework draws on multimodal semiotics and builds on the theoretical groundwork established in the preceding chapters. In particular, the guiding questions proposed by Packard et al. (45–46) inform the analysis. These include questions about the underlying syntax and affordances of each mode and how they are realized in the texts, as well as an examination of the relative prominence of each mode and its narrativity. Also, deviations from genre norms will be addressed, as well as specific intermodal relations that emerge, such as extension, parallelism, contrast, part–whole, or narrative relations (Packard et al. 60–61). These questions are addressed with a consistent focus on their narratological implications.

Analytically, this chapter adopts a comparative, text-based approach, reading *DWK* and *ATD* alongside one another in order to examine both shared conventions and significant differences. The study is firmly grounded in narratology, as well as multimodality, investigating classical narratological concepts such as focalization, narrative voice, and subjectivity in a multimodal context, while assuming that narrative meaning arises through the interaction of modes rather than within each mode independently. At the same time, the analysis remains genre-sensitive, consistently measuring its findings against the expectations of diary narration. Finally, the approach is qualitative and interpretive, as it foregrounds the narratological implications of multimodality in the novels under analysis.

6.1 The Fiction of Factuality and Intimacy in Multimodal Diary Narration

As discussed in Chapter 3.1, a supposed intimacy and factuality is a key presumption in diary narration, and this effect is significantly shaped within a multimodal context. Various visual elements contribute to a fiction of factuality (cf. Kellner), e.g. the design of the cover, the type of paper, or the choice of font, all of which work to evoke the appearance and feel of a real-life diary. Before turning to the individual analyses of the two works, the following section will first examine how this fiction of factuality is constructed and sustained within each text.

Noticeably, *DWK* uses lined paper and a font that imitates handwriting, and the cover looks like the diary that is described by the diarist himself. *ATD*, in contrast, does not use lined paper and the font of the running text resembles typed letters on a computer. As opposed to *DWK*, the cartoons in *ATD* seem to have been drawn on, partially, lined paper, which visually imitates pages torn out of a notebook and then taped into the diary, as can be assumed by the replication of tear marks and strips of tape on the edges. The implication here cannot be, however, that Junior uses a computer to note his thoughts and prints out pages with enough space in between the paragraphs to tape his cartoons in his diary. While the use of a handwritten font is an affordance of fictional diary novels, which can reinforce the fiction of factuality, its absence does not automatically produce the opposite effect, nor does it necessarily call the diary

framework into question. Rather, the fiction of physicality is a generic convention that readers are expected to accept. Diary novels remain prints of literary fiction meant to be reproduced, sold, and ultimately give the readers an entertaining experience. If a drawing on a piece of paper is taped in the diary, the reader has to accept that this is a print of a drawing with printed tape and not the original version. Thus, multimodal choices, on the one hand, can be crucial in sustaining the fiction of factuality. On the other hand, they should not be loaded with too much meaning beyond their functional and generic conventions.

Further, the assumption that both diarists, a twelve- and fourteen-year-old boy respectively, would be able to write a diary without mistakes, using literary devices and narratological strategies comparable to an adult author's abilities, seems unlikely.³⁷ Additionally, to imply a sense of timelessness³⁸, both novels abstain from using clear dates to head the diary entries. While Greg (*DWK*) uses months and weekdays, Junior (*ATD*) writes headings of chapters in a font different from the running text. Still, this intended vagueness leaves the quality of the novels as diary fiction unaffected, as specific dates would risk unveiling the fiction of factuality (cf. Field 82, 89).

Here, narratological principles such as the "principle of charity" (Walton 174–83) are important.³⁹ This means that certain fictive aspects are suppressed by the readers because they would make the fictive world seem paradox in an unpleasant way and would make imagining the storyworld more difficult (Packard et al. 76).⁴⁰ In the case of *ATD* and *DWK*, the reader willingly accept the fact that, despite not being writers, the diarists produce entertaining literary texts. They also have to accept the fact that the novels are fictional diaries, despite the lack of a usage of a handwritten font, lined paper, or concise dates.

As mentioned in Chapter 3.1, a diary novel is able to circumvent the fiction of factuality by employing a fictional editor, who wants to publish the memoirs of someone

³⁷ At least Junior is an avid reader, which may contribute to his writing skills. Greg, in contrast, shows little interest in books, making it less plausible that a twelve-year-old like him could master writing on that level, though, in *DWK*, the text mostly employs simple syntactical structures.

³⁸ This strategy no doubt is crucial to maximize the novels' commercial appeal.

³⁹ In German this is called the "Prinzip des wohlwollenden Verständnisses."

⁴⁰ This principle is exemplified by the fact that Othello speaks in polished verses, although he is not a poet, but a military commander (cf. Packard et al. 76).

special. Since this entity is present neither in *DWK* nor *ATD*, the reasoning for why they keep a diary and their motivation for what to do with it gains importance. This leads to the question of the diary's addressee and its fiction of intimacy. For both novels, a public audience is intended and addressed with the personal pronoun "you." In *DWK*, the audience is clearly the public, who in the future might take an interest in Greg's personal life. Although his mother made him start the diary, he claims that "The only reason I agreed to do this at all is because I figure later on when I'm rich and famous, I'll have better things to do than answer people's stupid questions all day long. So this book is gonna come in handy" (*DWK* 2). This statement is followed by a cartoon of Greg at a press conference, brushing off questions by curious reporters by referring to his diary.

In *ATD* there is no mention of an intent to publish Junior's memoirs, but, from the beginning, an implied reader is addressed by "you" several times throughout the novel and there are also rhetorical questions, as if Junior tries to communicate with an audience (e.g. "The typical human has thirty-two, right?," *ATD* 2, and "but let me tell you," *ATD* 4). Some clues, i.e. linguistic sensitivity, or the explanation of interpersonal relationships, indicate that this addressee could be a public audience, who knows nothing about Junior's situation. Although Junior and other characters, such as his best friend Rowdy, have a rather rough and vulgar way of talking⁴¹ – Junior claims that "Indians [...] love to talk dirty" (*ATD* 136) – there are no profane terms per se. Instead, family-friendly versions such as "fricking" (e.g. *ATD* 182) and "shoot" (*ATD* 227) are used, which seems an unlikely word choice for Junior. Also, he does note down his conversation with his basketball coach upon his visit to the hospital, claiming that "We told each other many stories. But I never repeat those stories. That night belongs to just me and my coach" (*ATD* 149), which at least suggests Junior expecting his notes to be read by someone else. He is sensitive about the addressee's values, which is why it can be assumed that this is a public as well. For instance, when he talks about enjoying masturbation, he anticipates that many might find this an inappropriate subject to discuss publicly and immediately justifies his declaration ("And maybe you're thinking, 'Well, you really shouldn't be talking about masturbation in public.'")

⁴¹ In a fight, Rowdy calls Junior a "retarded fag," for instance (*ATD* 52).

Well, tough, I'm going to talk about it because EVERYBODY does it. And EVERYBODY likes it," *ATD* 26, original emphasis). However, Junior is selective regarding which information he reveals to whom and under which circumstances: When he is interviewed before the basketball match, he refuses to answer the interview questions because he "wasn't about to share my feelings with the whole world" (*ATD* 183). It could be assumed that the diary gives Junior the power to control how his story is told, which serves as his reasoning for writing about his life.

If an intended public audience is assumed for *ATD* (for *DWK*, it is given), the reader can more easily believe the implied non-intimacy of the novel. Therefore, metaleptic instances, where obscene language is censored, interpersonal relationships are explained, or past events relevant for the story time but beyond the readers' knowledge are elaborated on, or even when the diarist introduces himself and his current living situation, can be accepted by the readers.

6.2 Case Study I: *Diary of a Wimpy Kid*

6.2.1 Greg as Diarist

In *Diary of a Wimpy Kid*, the protagonist Greg Heffley is explicit about his reasons for keeping his "journal" (1), namely his mother told him to record his feelings. He is sensitive about the label "diary," which he considers uncool, and fears might lead to him getting bullied at school. Although he initially finds the idea unappealing, Greg envisions a future in which he will be famous and too busy to explain his career to curious reporters, so he wants his memoirs ready for them. With this in mind, he addresses this future public with "you" throughout the novel. As a result, he also adopts a slightly narratorial style ("Now, I don't want to [...]," *DWK* 210) and carefully manages the information he shares, for instance by obscuring the fact that Rowley was forced to eat the cursed cheese, which would make him an outsider forever, if this information was known by others. However, the readers can easily infer what happened (*DWK* 210). This conscious awareness of his audience allows Greg to introduce new characters by explaining his relationships to them and characterizing them, without impairing the

fiction of intimacy, and it also enables him to build suspense, as in the moment when he opens Uncle Charlie's Christmas gift (*DWK* 123).

Greg writes entries fairly regularly, organizing them into month-based chapters, with individual entries marked not by exact dates but by weekdays – a device that avoids undermining the fiction of factuality, as discussed above. Although there are occasional gaps, he typically recounts recent events and tends to write only when something noteworthy has happened, usually soon after the event. Many entries refer to “today” (e.g. *DWK* 8) or “yesterday” (e.g. 21), and at one point he even writes “right now” (4) during a lesson at school, pondering on what to do “next period” (5). This is also the only instance in which the writing/drawing process itself is visualized, as there is a cartoon showing Greg at his desk with his diary and pen (see Figure 2). The fictional diary thus becomes part of the storyworld both visually and verbally, as it is also described as a red notebook labeled “diary” – something that annoys Greg, who insists on calling it a “journal” (*DWK* 1).

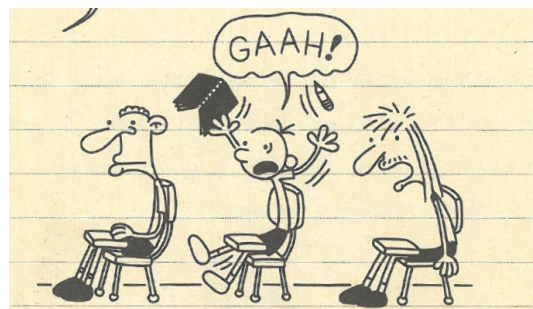


Figure 2: Greg shown in the process of journaling (*DWK* 4)

The length of Greg's entries varies considerably, ranging from less than a page to as many as twelve. Spanning the entire school year from September to June, the diary initially appears consistent, but reveals a clear decline in frequency of entries over time. On average, Greg writes just under seven entries per month, with September being the most productive with 15 entries, while June contains only one single entry. This decrease is not arbitrary but reflects the dynamics of the storyworld. After his fight with Rowley in spring, Greg loses both his main friend and, consequently, opportunities for noteworthy experiences to narrate. At the same time, this reduction underscores Greg's conception of his diary not as a means for reflective introspection, but as a curated memoir for his imagined future audience.

While Greg frequently comments on his reasons for writing, the act of drawing is never addressed within the narration. Yet the numerous cartoons, which carry significant narrative weight and advance the story, imply that Greg alternates between writing and sketching, even though this process is never mentioned. The only moment his interest in cartooning surfaces explicitly is when a school contest is announced to find a new cartoonist for the school paper. Greg's motivation, however, is not a love of drawing but the promise of popularity, since the previous cartoonist had become "a celebrity at our school and I wanted to get in on some of that kind of fame" (*DWK* 165). Still, as he collaborates with Rowley for this competition, Greg's enthusiasm and aptitude for drawing become evident, as he produces twenty strips "without breaking a sweat" (*DWK* 172) and even critiques Rowley's style, particularly his omission of noses.

6.2.2 Multimodal Meaning-Making

The relation between text and image in *DWK* can be described as an interdependent combination (cf. McCloud 155), as both modes provide essential information, and the narrative would be difficult to follow if one were removed (Reid and Serafini 43; Arend and Mahrt 929).⁴² Unlike many comics, there are few direct multimodal ties where the text explicitly points to the cartoon, for example when Greg writes "and Rowley came up to me and said – " (*DWK* 18), which is completed by a cartoon showing Rowley's embarrassing question, or when Greg introduces Fregley with "Like when he needs to go to the bathroom, he says – ," followed by Fregley screaming "JUICE!" (29). Such ties typically occur when speech is central, since the running text contains no direct dialogue and leaves that task to the cartoons. A rarer case of explicit linkage is a colon when Greg presents his haunted house plan: "Here was our final plan:" (*DWK* 54, emphasis added). Beyond these instances, text and image usually complement each other more loosely, as the text provides narrative framing while the cartoons render speech, emotion, and action. They also interact across the book, for example there is

⁴² This distinguishes *DWK* from the "girl version" *Dork Diaries*, a stylistically similar series by Rachel Renée Russell, following teenager Nikki Maxwell. In *Dork Diaries*, the images serve more illustrative purposes, which invite the readers to dwell longer in the situations the protagonist describes (Arend and Mahrt 930).

a contrast between Greg's reaction to Rowley's fart joke, where first he rolls his eyes (*DWK* 27), but later smiles contently (216), marking the development of their friendship (see Figure 3). In this way, meaning emerges not only from each mode individually but from their interplay across the diary.



Figure 3: Comparison of Greg's expression in *DWK* 27 and 216

Overall, the cartoons possess strong narrative power. While they occasionally repeat content, they mostly advance the story and add a temporal dimension to the spatial mode, since they frequently depict actions or conversations. Given the high number of cartoons in *DWK*, the overall narrative pace shifts toward a showing mode, creating both spatialized temporality and temporalized spatiality, and slightly conflating past and present. Because of their frequency, the cartoons also spatialize the temporal flow of the running text, pulling the reader into the narrated present even though the prose is in the past tense. The text itself is densely paced, leaving out many details, while the cartoons slow down the action and present it more synchronously.

To explore the multimodal interplay in greater detail, the verbal and visual modes will now be examined more closely in their specific functions. As mentioned above, the font of *DWK* mimics handwriting. Within the running text, spacing and paragraphing allow Greg to break down his narration into short, manageable segments, which are about as long as the images are high, thus, visually, both modes occupy the same amount of space. Verbal emphasis is achieved exclusively through capitalization, e.g. “this is a JOURNAL, not a diary” (*DWK* 1, original emphasis), while features such as italics or boldface, which would not fit the handwritten aesthetic, are omitted. In the cartoons, by contrast, written language appears in capitalized word balloons or as onomatopoetic sound effects like “RUMBLE” (*DWK* 136). Thus, while the running text simulates the personal act of handwriting, the cartoons adopt the conventions of

comic-strip lettering, signaling different communicative possibilities from the start. The running text in *DWK* is characterized by a sarcastic, colloquial style that conceptually leans toward orality (e.g. “stuff like that,” 81). More overtly oral expressions, however, appear only in word balloons in the cartoons (e.g. “Gee, thanks!” *DWK* 131), which reflects direct speech and thus contributes to the novel’s sense of immediacy.

Although the handwritten font in *DWK* strengthens the fiction of factuality, as discussed in Chapter 3.1, its realization also raises certain problems. When Gregs includes a note or a letter by a different character in his diary, thereby inducing a diegetic jump, he does not appear to insert the original artefacts, but rather reproduces self-made copies, maintaining the lined paper as a consistent background for the entire novel. For instance, he presents a questionnaire from his mother alongside his brother Rodrick’s answers, each distinguished by a different handwriting style, where the mother’s writing appears curly and playful, while Rodrick’s is rendered in sharp, oversized letters (*DWK* 42, see Figure 4). In this way, the font implicitly contributes to characterization, for example reinforcing Greg’s mother’s personality traits. This is supported by secondary literature, which points out that in the Arabic translation, the mother’s handwriting is not differentiated typographically, thereby weakening her characterization (Alós 261). Still, while the fonts serve this function, they also introduce inconsistencies. If Greg is the diarist, it seems strange for him to copy the distinctive handwriting of others. This oddity becomes even more prominent when he includes a cartoon of printed cards (still against the background of the lined paper!) he supposedly wrote for his family, where a computer-like typeface replaces the handwritten style (*DWK* 139, see Figure 4). A similar inconsistency occurs in the comic strip Greg submits for the school’s cartoonist contest, where the teacher’s editorial notes appear in a computerized font as well (*DWK* 178, see Figure 4). Finally, when Greg includes Rowley’s comics (*DWK* 170, see Figure 4), the handwriting is indistinguishable from his own, even though the drawing style differs, suggesting that Greg copies Rowley’s drawing style but not his handwriting. Since all cartoons appear directly on the lined paper background, it cannot be assumed that Greg sticks materials into the diary. Instead, it seems he

recreates notes, cartoons, printed texts, and posters, in which case multimodality actually destabilizes the fiction of factuality.

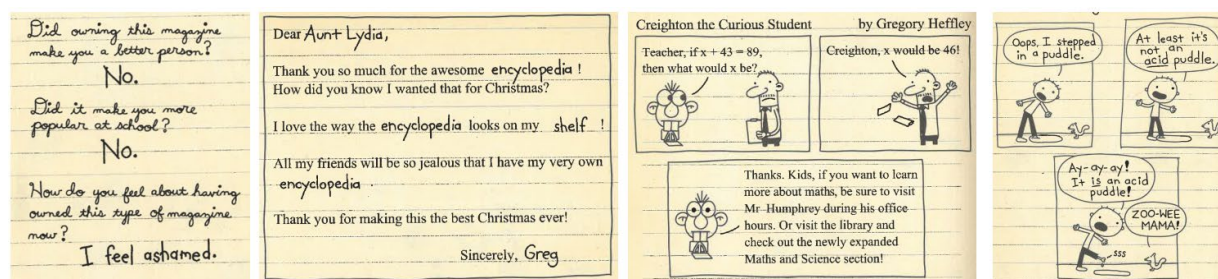


Figure 4: Stilted fonts in DWK 42, 139, 178, 170

Apart from characterization, another function of the running text is to move the narration along in a telling mode, recounting past events and not featuring direct speech.⁴³ The cartoons, by contrast, frequently use word balloons with direct dialogue or even abstractions of a conversation, as when Greg's father and a neighbor are shown "talking about football or something" (DWK 67), represented simply as "Blah blah blah [...]". This captures Greg's perspective and strengthens the fiction of factuality, since he could not reproduce the actual conversation. Overall, while the running text tells, the verbal elements in the cartoons mostly have a showing quality, directly relaying unmediated speech. Apart from that, written language in the cartoons is also used when posters or plans are represented, such as Greg's campaign for school treasurer (47, see Figure 5), employing Greg's regular handwritten font. Only twice is it used as annotation, once to label Greg's Halloween costume with "double baseball hats" (63, see Figure 5), and once to identify a small "drawing pin" (200). Written language also serves to clarify actions, as in the cartoon where the word "SHOVE" (114, see Figure 5) signals a bully pushing a kid to the ground. Beyond relaying dialogue, another important function of written language in the cartoons is the representation of sound, most often through onomatopoeia. Respective depictions are visually varied in size, shape, and font to reflect intensity and quality, for instance, the bold "VROOM" (17) of a loud vacuum cleaner, the jagged "SLAM" (25) of a surprisingly shut door, or the crashing "SPLASH" (74) when Greg's father empties a bucket of water on his son and Rowley (see Figure 6). Speech

⁴³ There is only one rather indirect instance of direct speech featured in the running text, namely a mediated "So I just said 'thanks' instead" (DWK 126).

balloons themselves are often modified to convey auditory qualities, for example, a zig-zag outline signals a whistle (“TWEET!” 83, see Figure 6), a drooping balloon mirrors Greg’s father’s menacing tone when emphasizing “friend” (37), and dotted outlines indicate whispering (103) (see Figure 7). Typography also contributes to vocal characterization, as in the pixelated font of a robot’s voice (148) or the wavering letters of Rowley’s exaggerated “Want to come over to my house and plaayyy” (18) (see Figure 7). Volume and emphasis are further conveyed by letter size, such as the diminishing text of the trees’ ever more quieting song (111, see Figure 7). Visual markers like zig-zag musical notes for Rodrick’s Heavy Metal music (36) or lines emanating from characters’ mouths to represent shouting (e.g. 187) also expand the soundscape of the diary (see Figure 8).

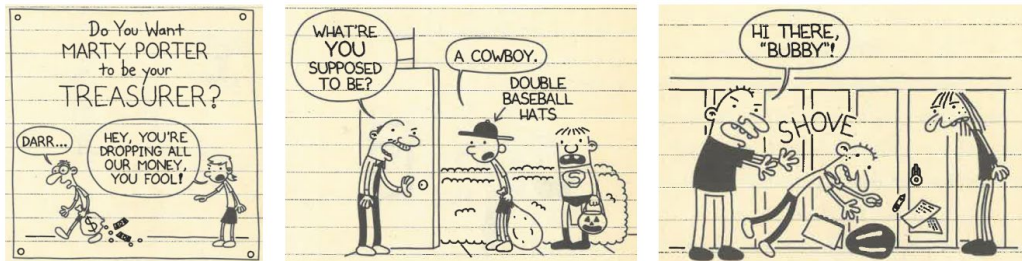


Figure 5: Verbal Written Language in the Cartoons (DWK 47, 63, 114)



Figure 6: Onomatopoeia in DWK (17, 25, 74, 83)



Figure 7: Sound quality in DWK (37, 103, 18, 148, 111)



Figure 8: Representations of sound (*DWK* 36, 187)

Although onomatopoeia and paralinguistics may appear as minor features, they play a crucial role in shaping the series' tone. Translation studies highlight this importance, in the Arabic edition, for example, the omission of such elements significantly reduces both Greg's expressiveness and humor (cf. Alós; Al-Jabri).⁴⁴ Similar findings emerge in research on the translation of ideophones into Italian (cf. Pischedda) and expressive language into German, Croatian, and Serbian (cf. Prodanović Stankić and Begonja), underscoring that sound effects are fundamental to the multimodality of the text. In fact, these small markers actively contribute to Greg's narratorial stance, by revealing his subjective experience of how events sound and feel to him.

The visuals in *DWK* maintain a consistent, cartoony style and always stretch across the page width, interrupting the text and creating a clear reading path. They are tightly interwoven with the running text, advancing the narration while adding humor. Most panels adopt a long shot (e.g. *DWK* 56) or medium-long shot (e.g. 42), often show figures aligned on the same axis, and perspective shifts are rare (see Figure 9). While most cartoons depict a single scene, some appear in short sequences, forming mini comics (e.g. 168, see Figure 9). In most cartoons, the process of closure also plays an important role, as objects, such as the table in Figure 9, are shown only partly and are completed mentally by the reader. Although the verbal narration quantitatively slightly outweighs the visual narration, there is at least one cartoon on every page.

⁴⁴ Both Alós and Al-Jabri offer valuable insights into the Arabic translation, noting that in Arab countries greater emphasis is placed on a book's educational rather than its entertainment value. They also highlight several ideologically motivated changes affecting both modes, so besides changing certain expressions, depictions of Christmas decorations are altered, and cartoons of women in bathing suits are modified.

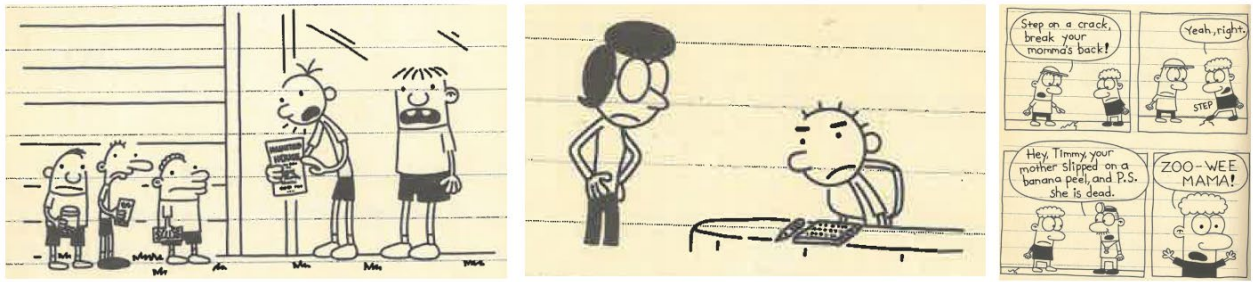


Figure 9: Long shot (DWK 56), medium-long shot (42) and mini-comic (168)

Generally, the featured cartoons serve multiple narrative functions beyond illustrating the text. They typically depict present events, but also show imagined futures, past memories, alternative scenarios, wishes, or visual artifacts such as posters, plans, and letters. Often, they elaborate on the running text by visualizing possible situations and frequently they advance the plot independently of the text, showing the action rather than narrating it. Importantly, the cartoons convey emotional nuance and exaggeration, shown in Figure 10, as in the scene where Jason Brill tries to sit next to Greg. While the text merely states that Greg prevented it, the cartoon dramatizes Greg's frantic refusal and gives the reason why he would not want to sit next to Jason by exaggerating his rowdy appearance (DWK 5). Further, through the long shot-perspective, the cartoons often add information about a situation (for example showing Greg's Dad wearing a cap on Saturday morning, implying that he had already been outside, 17) or provide details about a character (for example showing Greg's messy room while he sleeps in during the holidays and his alarm clock revealing that it already is 1:20pm, 16, see Figure 10). Overall, the cartoons mirror the sarcastic undertone of the running text and, as Arend and Mahrt note, they often serve to punctuate verbal narration with punchlines or attitudes, significantly contributing to the overall comedic stance of the novel.



Figure 10: Emotional nuance (DWK 5) and indirect characterization (16)

The cartoons in *DWK* employ a typically iconic vocabulary and rely on the use of visual metaphors. One icon, for example a small cloud, can be interpreted in various ways, depending on the context. Thus, this cloud could indicate jumping, slipping, shaking, or exhaust fumes (e.g. *DWK* 44, 106, 77, 76, respectively), as Figure 11 shows, illustrating the complexity of visual metaphors. This way, they are also able to represent absences or missing elements, such as Greg's blank thought balloon when he cannot think of ideas (*DWK* 138) or four lines indicating the missing Cheese (213, see Figure 12). As far as the depiction of the characters is concerned, they are abstracted yet easily identifiable, each with a singular appearance, even minor figures who appear only once. Key features are often simplified, like Greg's hair consisting of just three strands. While many characters have different hairstyles or facial shapes, other physical traits are also simplified and there are visual ties between certain characters, distinguishing them from others, for example, the Heffleys have round faces, Rowley a rectangular one, and Fregley a pear-shaped one. Outfits are similarly consistent, changing only for narrative reasons, such as outdoor winter clothing. The facial features are minimal, as the eyes are small dots and the mouths simple lines, which can signal mood with minimal means, e.g. a frown or happy smile, where the line depicting the mouth is bending down or up. Usually, Greg is shown with a droopy mouth, suggesting boredom or indifference. If rarely depicted physical details, such as teeth, are shown, this trait is emphasized, such as Rowley's buck teeth. Motion is conveyed visually via lines indicating movement or shaking. Similarly, the passing of time is also expressed relying on visual cues, for example, when a sequence of actions is shown in one cartoon, as seen in Figure 13. In this example, Rowley and Greg are in Mr. Winsky's office, and Greg interrupts Rowley by pinching him under the desk (153). Here, the co-presence of both the neutral beginning of Rowley's "But we would miss [algebra]," as well as his pained face and expression of surprise ("YAHOOEY!"), express the passage of time by showing two consecutive moments simultaneously.

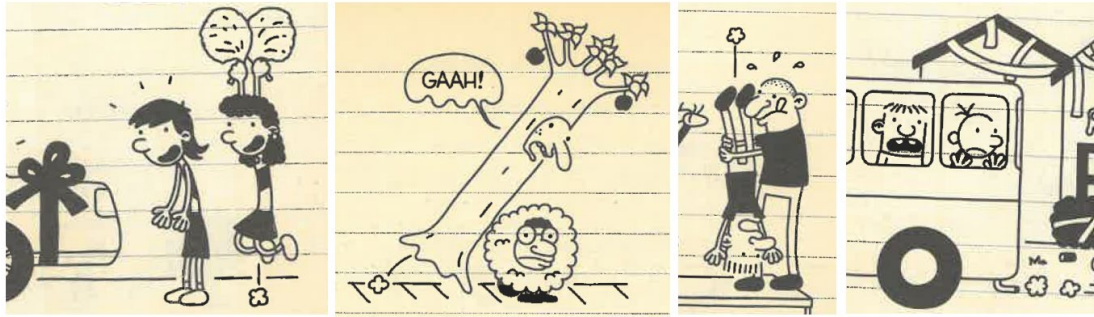


Figure 11: Semantic range of the cloud as visual metaphor (DWK 44, 106, 77, 76)

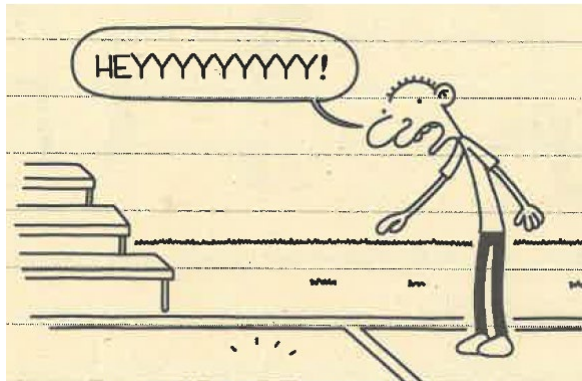


Figure 12: Showing what is not there (DWK 213)

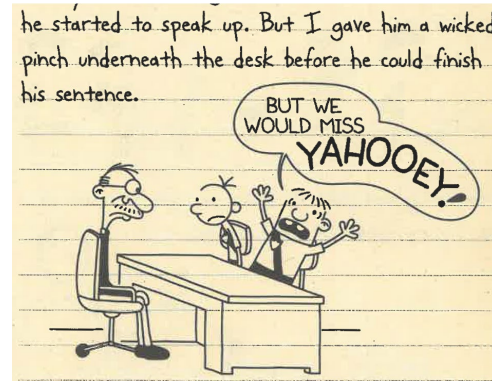


Figure 13: Multiscenic cartoon (DWK 153)

The iconic vocabulary of the cartoons plays a crucial role in conveying emotions and visual metaphors, particularly because Greg rarely articulates his feelings verbally. Emotions are instead expressed through visual cues, like gestures (face palm), exaggerated facial expressions, or visual metaphors (sweat beads). Given the exaggeration of emotions, as is typical for iconic depictions, it remains open whether Greg actually facepalms (e.g. as a reaction to Rowley's confession, DWK 35, see Figure 14), or if this is a metaphorical representation of his feelings. Given the showing quality of the pictorial mode, the cartoons amplify emotions, for example, when Greg and Rowley walk home on Halloween: their wide eyes, gritted teeth, and sweat beads convey fear beyond the verbal description of "Snake Road is scary" (73, see Figure 14). Further, the cartoons can also nuance emotions. For example, Figure 15 shows the Heffley family exhibiting different levels of anger after the school play disaster (113): Greg's mother has a small tornado above her head, Rodrick shows inward-pointing eyebrows, and his father appears comparatively less angry without these features. Usually, Greg's visual appearance reflects his general emotional neutrality, occasionally portraying personal affections, such as learning that Rowley becomes

the new school newspaper cartoonist, where he reacts with an open mouth and three exclamation marks above his head (204, see Figure 14). Given the overall consistency of the visual cues, the intersemiotic relations formed can be described as parallel and narrative in terms of Packard et. al (61).



Figure 14: Metaphorical expressions of emotions (DWK 35, 73, 204)

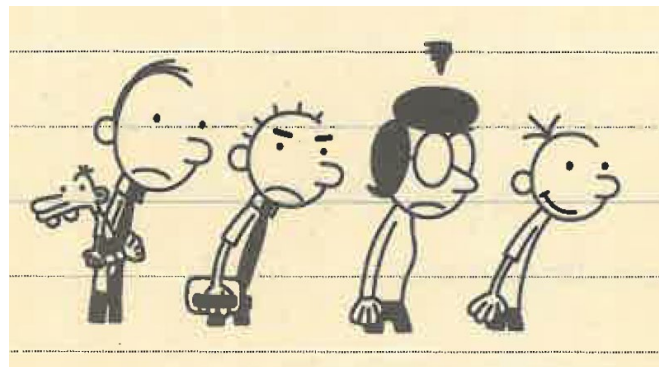


Figure 15: Different levels of anger (DWK 113)

To conclude the discussion of visual elements in *DWK*, it is noteworthy that while the cartoony style emphasizes universality and Greg does not age throughout the series,⁴⁵ the visual mode can also anchor the story in a specific time. Unlike the written text, which could simply mention watching TV, the cartoons specify the era through depicted objects, like a 2007-style television (DWK 23), radio (91), video recorder (108), and phone (60), as shown in Figure 16. This illustrates how visual details can subtly date the narrative, even in an otherwise timeless series.⁴⁶

⁴⁵ Although the most recent volume, *Partypooper* (as of March 2026), features Greg's birthday, he does not age over the course of the series, nor do the other recurring characters. In this regard, the iconic cartoony style proves particularly advantageous, as an abstract character can continue indefinitely without undergoing realistic aging. This was also remarked by Jeff Kinney in an interview with "The Guardian", where he said that Greg is fundamentally a cartoon character and as such can go on for a long time (Heritage).

⁴⁶ Some researchers have claimed that still, *DWK* is timeless in that it abstains from references to current trends in popular culture or worldly occurrences. This distinguishes it from, for example, *Dork Diaries*, which is considered less timeless (Rickard 6).



Figure 16: Outdated technology in *DWK* (23, 91, 108, 60)

6.2.3 Constructing Narrative Stance

Having established the distinct roles and affordances of written language and visual elements in *DWK*, it becomes possible to consider how these modes work together to shape the reader's experience of the narrative. The interplay between text and image does not merely present the story, it also influences how the narrator's voice is perceived, how events are focalized, and how the diarist's subjectivity is conveyed.

Greg Heffley is the diarist and thus the narrator of *DWK*. His stance is highly self-centered, marked by sarcasm, humor, and a lack of reflection.⁴⁷ In his entries, he does not attempt introspection but rather entertainment. Given the overall showing mode of the text, which contrasts monomodal diary novel narration, the focus automatically lies on the story rather than the discourse time. As opposed to generic diary narration, multimodality can shift the focus away from reflection, to telling a story to a clearly envisioned audience. Greg consistently addresses this imagined audience and aims at entertaining them. The cartoons foreground humor and action over reflection, visually enacting Greg's tendency to avoid self-analysis. The showing quality that is emphasized by the text's multimodality shifts the focus on the received present of past actions, drawing the readers into these situations more effectively.

Greg's voice is ironic and he often downplays events or shifts the blame onto someone else. The visual mode, however, reveals what he fails to admit in his writing. For instance, when Greg insists that he is unaffected by Rowley's accident, the accompanying cartoon shows his shock over the situation (*DWK* 140, see Figure 17).

⁴⁷ It seems worth noting here that Greg is a twelve-year-old boy, who might not yet have developed profound skills of self-reflection. Generally, the novel's humor is certainly shaped by Greg's obliviousness toward the consequences of his own actions.

The coexistence of two modes produces a tension between the narrating I, who tries to control the narrative voice, and the experiencing I, who cannot be fully hidden because the images expose his reactions. Since Greg is the one who also produces these images, this could support his inability to reflect, and instead channeling his emotions in a more immediate way, i.e. through images. This way, Greg's lack of emotional reflection in the running text is often counterbalanced by the visual mode. This creates a striking contrast. For instance, while Greg verbally does not mention the possibility that Rowley is angry with him, the accompanying cartoon clearly shows Rowley with an angered facial expression (184, see Figure 17), indicating that Greg does perceive his friend's emotions. Thus, even when his narration downplays or distorts emotions, the images reveal a more accurate version of events. On one level, Greg consistently portrays the emotions of others, yet at the same time he appears oblivious and indifferent to them. In the haunted house scene, for example, he ignores the fear of the first child and instead worries only about "all the money we were losing with this kid clogging up the Hall of Screams" (57, see Figure 17). In such moments, multimodality highlights Greg's conflicted, or maybe even lacking, sense of empathy, exposing gaps between what he writes, what he shows, and what he seems to feel.

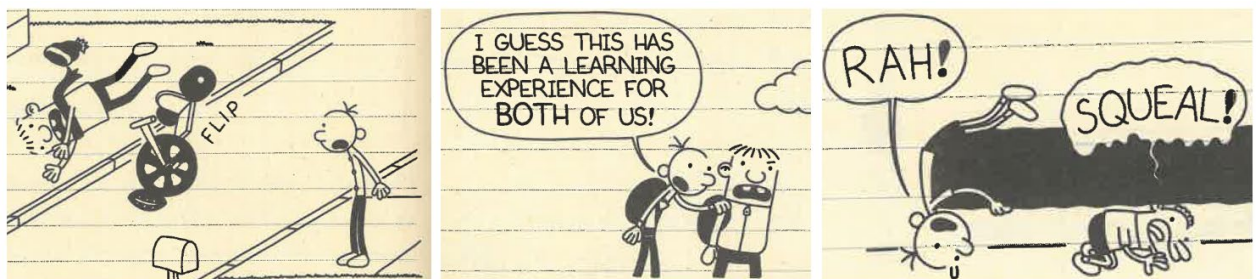


Figure 17: Revealing emotional expressions in the cartoons (DWK 140, 184, 57)

It is exactly this tension arising from the interplay of the modes that shapes the novel's humorous tone. Sarcasm in the text often contrasts with exaggerated visualizations, producing irony. To mention one example of this: When Greg confesses that he was the one who got Rowley into trouble, he reframes the situation as a shared learning opportunity, even advising Rowley to be more cautious about whom he lends his hat to. He then wonders why his "message didn't seem to be getting through to Rowley" (DWK 184). When his friend subsequently ignores him, Greg dismisses him as a "jerk" (DWK 189), but also says that he "tried to break the ice with [Rowley] today, [...]. But

even THAT didn't seem to work" (189, original emphasis). As Figure 18 reveals, "THAT" means saying "Hi, pal!" (DWK 189) and hitting Rowley with a snowball. This act of aggression rather than reconciliation underscores the irony of Greg's claim to be reaching out and highlighting the gap between his verbal narration and the visual evidence. This contrast is realized through the multimodal context, showing Greg's unreflected obliviousness, while highlighting his unreliability as narrator, as well as amplifying the comedic effect. Greg has a cool, sarcastic voice in the text, while the cartoons involuntarily show his vulnerability. The visuals often reveal that Greg is not the good guy he portrays himself as. Thus, multimodality and its contradictions of text and cartoon complicates the first-person unreliability (Rickard 53). Greg is a prominent and easily recognizable example of unreliable narrators in children's literature. His unreliability stems not from deliberate deception but from self-absorption and lack of self-awareness. He consistently downplays his flaws, exaggerates others', and omits inconvenient truths. Yet the images often serve to expose Greg's lack of empathy or self-awareness. In the above-mentioned episode, when Greg finally admits to Rowley that he let him take the blame for his own wrongdoings, he neither apologizes nor shows any sign of guilt. Instead, he assumes Rowley is simply "tired" (DWK 184) when he refuses to hang out afterward. Yet the cartoon accompanying this scene clearly depicts Rowley with an angry expression. Either Greg's narration fails to acknowledge what he perceives, undermining the fiction of factuality, or he is more capable of expressing emotions visually than verbally. In both cases, the image reveals an awareness that his written words suppress.

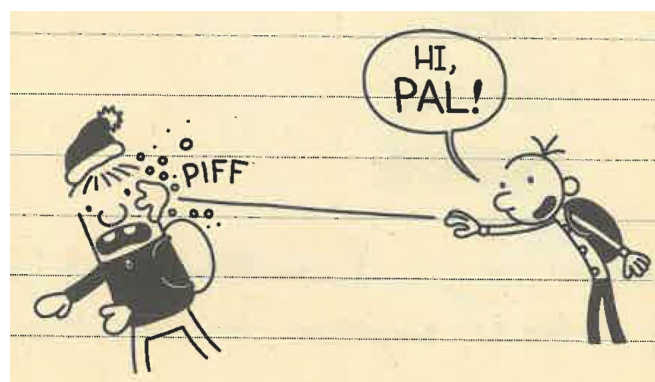


Figure 18: Greg's attempt at reconciliation (DWK 189)

Further, Greg, as the autodiegetic diarist, has the narrative dominance in the novel (cf. Hui, “Narrative Voice”). Other characters are primarily filtered through Greg’s perspective, and when he, as he rarely does, includes their speech indirectly in the running text, it usually serves to belittle them. For instance, he quotes Rowley’s need for a “potty break” (*DWK* 68) in order to highlight Rowley’s childishness (Hui, “Narrative Voice” 404). In this way, Greg carefully controls the information the reader receives, which makes his narration particularly prone to unreliability. However, as Hui notes, the illustrations open up space for resistance to this dominance, as they allow other characters’ voices and expressions to emerge more directly, offering glimpses of their authentic personalities. These visual moments serve as a counterbalance to Greg’s potentially skewed reporting, granting the narrative an immediacy and polyphony that is largely absent from the written text (“Narrative Voice”). Further, multimodality provides possibilities for diegetic jumps, such as when Greg includes Fregley’s letter or his mother’s questionnaire, which briefly interrupts his autodiegetic voice and authority over the narrative.

Narratologically, focalization in *DWK* is not stable. The written text is consistently homodiegetic, as Greg is both the narrator and protagonist. Yet the images complicate this, at times visualizing subjective impressions of other characters, at other times providing an external, non-subjective, vantage point. One example that shows Greg’s subjective focalization, is when he describes Snake Road on Halloween as scary and the image amplifies his fear: Greg and Rowley are shown with wide eyes, clenched teeth, and sweat beads (*DWK* 73, see Figure 14). This is less a realistic depiction and more a subjective visual metaphor for fear. Similarly, when Rowley laughs at a fart joke, Greg supposedly rolls his eyes (27, see Figure 3). Whether Greg actually performs the gesture or merely imagines it remains ambiguous, but it is possible that the cartoon visualizes not just actions, but inner states of mind (i.e. internal focalization). Whether the cartoons show a subjective impression or an intersubjectively valid scene, almost all of them adopt an external, third-person view. There are, however, two notable exceptions where the images shift to a first-person perspective, which are shown in Figure 19. The first shift occurs when Greg peeks through the curtain before the school play, revealing Rodrick in the audience (106).

The second shows Greg receiving Fregley's letter under the bathroom door, with his finger touching the booger stuck to the note (194). In both cases, the perspective shift underscores Greg's immediacy of experience and heightens the subjective involvement of the reader, contributing to a suspension arch. More often, however, subjectivity in the cartoons is conveyed through exaggeration rather than point of view. For example, it seems unlikely that Greg's younger brother could realistically be seen standing on a stool behind his parents, loudly shouting Greg's embarrassing nickname during the school play (108), as seen in Figure 20. Instead, the image reflects Greg's perception of the moment, amplifying his sense of public humiliation rather than depicting the literal arrangement of the scene. One striking instance of subjectivity even originates from another character, namely Mrs. Irvine, and is also shown in Figure 20. She fails to recognize Greg from a distance, the cartoon shows him only as a black silhouette, indistinguishable from the other children (181). Here the panel adopts her perspective but still renders it in an externalized third-person view, showing her from behind while simultaneously visualizing what she perceives. Thus, focalization is more flexible and hybrid in this multimodal context. Multimodality also opens up perspectives beyond Greg's direct experience. While he usually illustrates scenes in which he is personally involved, there are several instances where the cartoons depict events he could not have witnessed. For example, Manny showing Rodrick's magazine to his kindergarten classmates (41) or the principal yelling at Rowley (180), are moments Greg was clearly not present for, yet they are still visually represented (see Figure 20). In other cases, Greg omits himself from the cartoon even though he might plausibly have been present, such as when his mother dances to Rodrick's music (31). These images suggest that Greg occasionally reconstructs scenes second-hand, either from what he has been told or from his own imagination, extending the narrative beyond his immediate perspective. Often, he visualizes his imaginations of a possible or impossible future, for instance, when he presents himself as being very muscular in P.E. (87, see Figure 20). Thus, in a manner typical of diary narration, most visualizations are subjective, despite showing Greg from a strictly consistent outside view.

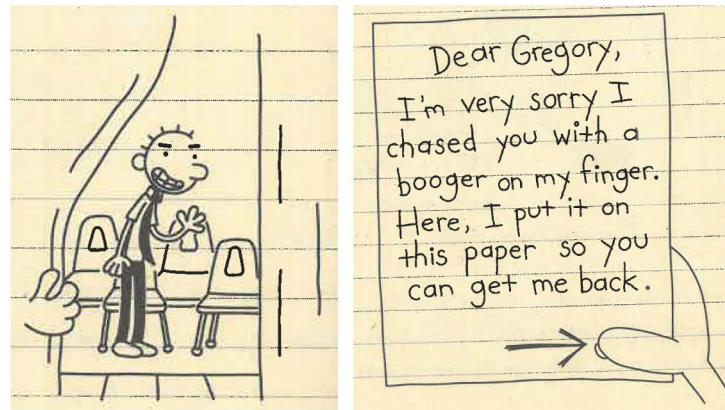


Figure 19: Focalization shifts (DWK 106, 194)

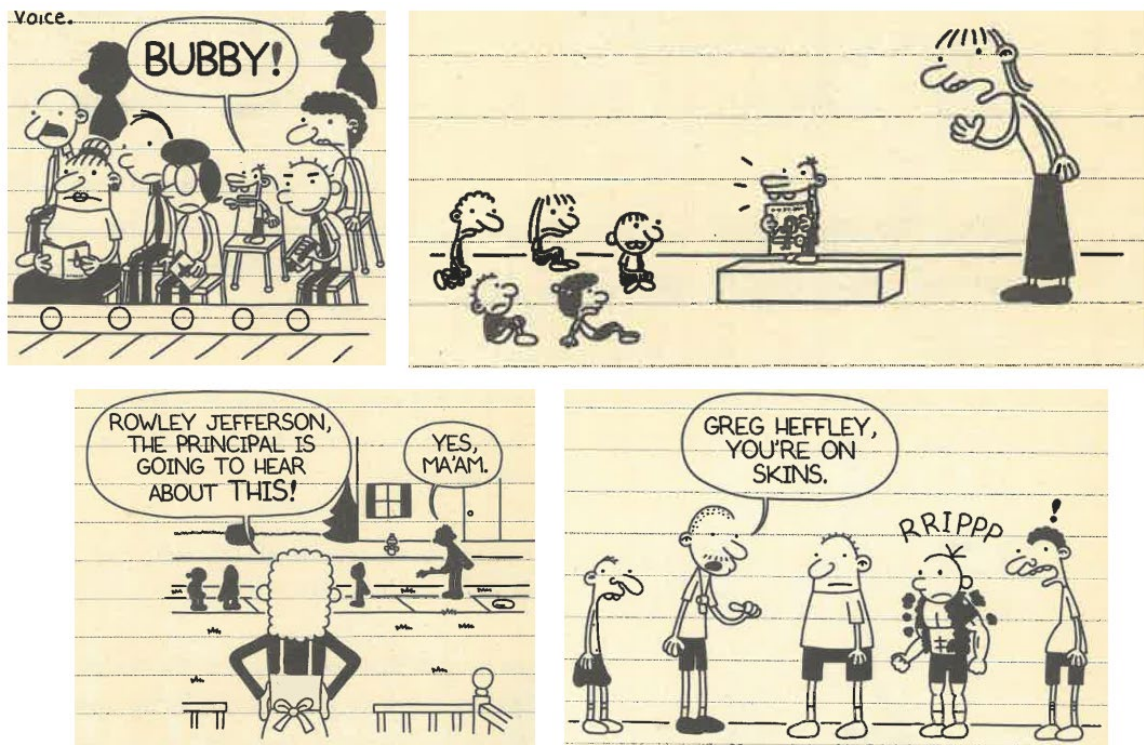


Figure 20: Subjectivity in DWK (108, 41, 181, 87)

Where focalization is split is not between Greg and other characters, but between the narrating and experiencing I. The running text typically focalizes the story through the narrating Greg, whereas the images usually focalize through the experiencing Greg. As a result, readers are offered both perspectives. Since the visual mode tends to be received more immediately, the connection to the experiencing Greg is often stronger than to the narrating voice. This dual focalization creates a hybrid position for the reader: They share the narrating I's foreknowledge while simultaneously empathizing with the frequently oblivious experiencing I. A striking example occurs when Greg sneaks into the living room at night to listen to his older brother's heavy metal CD

marked with a parental advisory sticker (*DWK* 36). He intends to use headphones and turns the volume to maximum. The cartoon, however, already reveals that the headphones are not plugged in. Through visual parallelism (cf. Packard et al. 52–58), the reader can also infer this because Rodrick listening with headphones is depicted without musical notes, whereas Greg is surrounded by them, as shown by Figure 21. Thus, both the narrating I and the reader recognize the impending consequences, while the experiencing I, smiling in front of the stereo, remains unaware. In this way, the relationship of text and image positions the reader as a more perceptive observer of the story than the experiencing Greg himself.

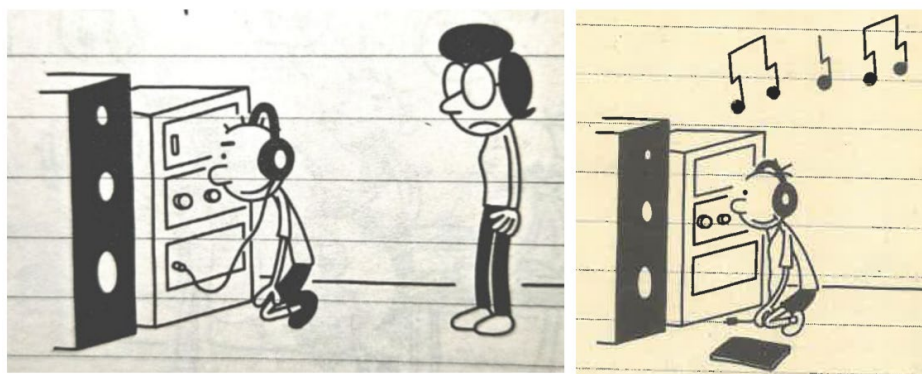


Figure 21: Visual parallelism and contrast in *DWK* (32, 36)

6.2.4 Findings

Ultimately, *DWK* departs from the conventions of monomodal diary narration by integrating the visual mode in a way that significantly shapes its narratology. Unlike traditional diary novels, which often foreground introspection and reflection, Greg's narration is predominantly oriented toward recounting events. This emphasis on a telling mode, however, is not confined to the verbal narration, but is equally advanced by the visual mode. The cartoons do not function as sites of reflection but rather dramatize actions, dialogue, and comic situations. As a result, the diary reads less as a record of inner development and more as a sequence of anecdotes centered on external events. In this context, it is particularly striking that the act of drawing is never explicitly motivated within the narrative. Greg's affinity for cartooning, his skill, and his decision to narrate visually remain unexplained. At the same time, this absence is narratologically meaningful, exactly because the cartoons are well suited to narrate

events but less apt to convey introspection. Multimodality thus reinforces Greg's limited manner of self-reflection. By consistently presenting himself from an external perspective, using highly uniform perspectives and stances, the cartoons underscore his lack of inner complexity.

Further, a hybrid focalization between the narrating and experiencing Greg is enabled in the multimodal context of the novel. The written text reflects the perspective of the narrating I, who recounts past events with varying degrees of reliability, whereas the images present the experiencing I, capturing Greg in the immediacy of situations. Readers are therefore positioned at the intersection of these two viewpoints. They possess the retrospective knowledge of the narrator while simultaneously empathizing with the often clueless or short-sighted protagonist. This hybrid focalization invites readers to evaluate Greg's behavior critically, as they perceive both what he reports and what he fails to grasp.

At the same time, the interplay between word and image generates much of the humor and sarcasm that characterize the series. Frequently, there is a dissonance between what Greg writes and what the cartoons depict, which produces comic irony at his expense. For instance, while the written text often presents Greg as oblivious, defensive, or self-justifying, the drawings simultaneously reveal emotions, reactions, or consequences that undermine his claims. This tension between the modes is not incidental but structurally embedded. While the running text establishes Greg's narrative authority, the images destabilize it by exposing contradictions, exaggerations, or overlooked details. Humor thus emerges not only from Greg's sarcastic tone in the narration but also from the multimodal gaps that force the reader to recognize what Greg himself cannot or will not admit.

6.3 Case Study II: *The Absolutely True Diary of a Part-Time Indian*

6.3.1 Junior as Diarist

Arnold Spirit Jr.'s diary departs in several respects from the conventions typically associated with diary fiction and is identifiable as such primarily through its title. The entries do not begin with the typical "Dear Diary" or are organized by dates. Instead, they are structured more like chapters in a novel, each introduced by a title prominently displayed on a new page. This formal arrangement distances the text from the characteristics of diary novels. As mentioned above, the modes also do not try to adopt the theme of diary novels by using a handwritten font or referencing the diary or the act of writing it. Overall, *ATD* resembles a first-person narrative more than a traditional diary novel.

The writing process in *ATD* is never mentioned explicitly, the reader does not learn when or where Junior writes his entries and the intervals have to be deduced implicitly by clues in the text. Junior's diary has 29 entries, which are more structured like chapters, spanning from the beginning of the school year and his decision to attend a school outside of the reservation, to the beginning of the summer break and him reconciling with his friend. Given the fact that he does not ever mention a date, nor when, where or why he writes in his diary, assumptions have to be made indirectly from clues in the text. It can be assumed that the entries are never made on the same day as the events occur, as the events retold in the entries often span over a few days. Whether he writes these long entries in one session or takes several writing attempts to finish a chapter remains unclear. This way, he gives a detailed account of his first day at his new school, only to say that he did not speak for six days, and elaborates on what happens on the seventh day of school in great detail again. This entry is followed by a two-page story about a conversation he had with Rowdy four years prior, which is in turn followed by an entry about Halloween. These intervals continue throughout the whole diary and, despite the long breaks in between,⁴⁸ Junior gives detailed accounts of the beginning of these time spans. There is only one instance of him writing an entry

⁴⁸ The long intervals between the entries, especially after the deaths of family members or close friends, might help characterize Junior, as he potentially needs some distance to these tragic events before he is able to verbalize them.

on the same day, which he starts off with “Today at school [...]” (*ATD* 130) and only encompasses two pages. The entry encompassing the longest time span is a chapter called “Reindeer Games” (135-49). In this chapter, Junior’s father convinces him to join the basketball team, and the first match against Wellpinit two weeks later is described as well. During this match Junior is knocked out in the third quarter and has trouble remembering the rest of the events, which is why he retells them as “secondhand information” (146). At the end of the chapter, his coach visits him in the hospital. Sometimes, Junior does not reflect on events from the recent past, but from four years ago, giving a verbatim account of conversations he had with Rowdy, intertwined with stories his father told him further in the past. These accounts are usually prompted by similarities to the present situation. Similarly, Junior uses the foreshadowing of events over extremely long time stretches.⁴⁹ In other cases, he gives detailed accounts of being physically assaulted, with a level of situational awareness, which seems implausible. All in all, Junior’s extended entries, which alternate between detailed episodes and more cursory overviews of events, undermine the conventions of the diary novel and the fiction of factuality, aligning the text more closely with a non-diary novel featuring an autodiegetic narrator.

In contrast to the writing process, Junior states an explicit motivation for drawing cartoons in his diary. He repeatedly reflects on the reasons behind his artistic expression and articulates its personal and social functions. Because of his physical disabilities and resulting appearance, he is stigmatized as a “retard” (*ATD* 4) on the reservation and subjected to frequent violence, which leads him to withdraw from social life and instead immerse himself in reading and drawing. He explains that he prefers cartoons because “words are too unpredictable [and] limited” (5), whereas images are universally intelligible and thus give him a more reliable medium of communication. His

⁴⁹ There are two particularly striking instances of foreshadowing that undermine the fiction of factuality. The first occurs in the entry written only a week after Grandmother Spirit’s death, where Junior recounts the later fate of the drunk driver responsible (“After he got out, Gerald moved to a reservation in California and nobody ever saw him again,” *ATD* 158). Since this information would only have become available a year and a half after the accident, it clearly exceeds the diarist’s knowledge at the purported moment of writing. The second instance appears after the Winter Formal, when Roger drives Junior home and the narration notes, “And he drove me home plenty of other nights, too” (129). Here, too, the statement presupposes knowledge of future events, further destabilizing the temporal immediacy and authenticity that diary narration conventionally conveys.

main goal is “to talk to the world” (6) and to be heard, with the hope of one day becoming a famous artist. In this sense, he perceives his drawings as his “only real chance to escape the reservation” (6). For Junior, cartoons serve both aesthetic and relational purposes, as he claims that he draws in order to understand and make fun of the world as well as to honor friends and family (95). Moreover, his drawings are presented as an act of care, as when he states that he sketches in order to make Rowdy happy and to “give him other worlds to live inside” (23). As Garić notes, the creation of alternative, better worlds constitutes a timeless function of storytelling, and Junior’s cartoons participate in this tradition. Apart from this social function, Junior also sees drawing as central to his identity and self-worth, as he claims that “I’m a cartoonist. And I think I’m pretty good at it, too” (7). Junior never makes comparable statements about writing, although he clearly has a talent for this form of artistic expression as well, given his sophisticated narrative voice throughout the novel. However, he positions himself as a reader, which shows his general tendency toward creative ways of expressing internal states.

The importance of artistic self-expression with regard to Junior’s development and its therapeutic purposes have been highlighted by Garić as well. Junior’s drawings serve as a means to heal and channel his emotions, “emptying their disturbing contents into caricatures and thus reducing their strength, their reality” (Garić 194). This way, his cartoons are a way “of nourishing what he loves” (194). Garić also highlights the role of “art as an agent of survival” (194), which becomes apparent upon Junior describing his cartoons as “tiny little lifeboats” in a world of “a series of broken dams and floods” (*ATD* 6). In tough times, after his grandmother and Eugene pass, Junior mentions drawing a lot (“And I drew and drew cartoons,” 171), but also finding solace in reading, saying stories helped him during that time (“I needed books. I wanted books,” *ATD* 171). He also compiles lists of things that make him happy and keeps “drawing cartoons of things that made me angry. [...] It became my grieving ceremony” (178), making drawing and redrawing a crucial part of his grieving process. However, these cartoons are not included in the diary.

Although the overall process of drawing is not explicitly mentioned, some aspects are revealed in the narration. The reader thus learns that Junior draws in his room, for

example. However, the process becomes tangible, for example in his drawing of Rowdy, which is shown in Figure 22. Here, Junior exhibits skills beyond cartooning and draws his friend with fine lines, showing clear details. Rowdy is lying on his belly, with his hands supporting his head, reading a book. There is, however, an angry cartoony face drawn over the sketch of the more detailed face, revealing sharp teeth and zig-zag lines and sweat beads emanating from the top of his head with a speech bubble saying “What’re you drawing??” (*ATD* 23). The drawing is also commented by Junior: “He hates it when I draw him! Never lets me finish” (23, original emphasis). In an interview with Ellen Forney, the novel’s cartoonist, attached at the back of *ATD*, she describes her reasoning for using different drawing styles. She states that the “scribbled-looking illustrations and comics” are jot-downs of Junior’s thoughts, whereas the “slightly more realistic cartoons, like the annotated portraits of his family” show a more time consuming drawing process and lastly “the penciled portraits” (*ATD* “Interview” 240) suggest a long time spent with the person Junior draws or using a portrait or photograph as a reference.⁵⁰ This copying of a photograph is explicitly mentioned by Junior when he draws his sister’s new trailer, which he drew “From a photo she sent of her ‘gorgeous new place’” (*ATD* 134).

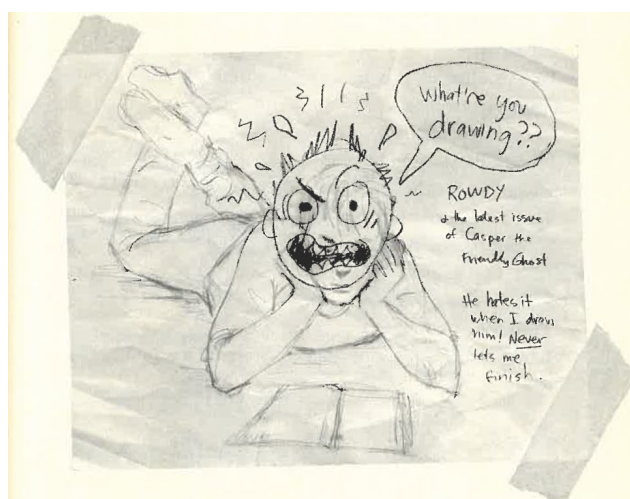


Figure 22: Sketch of Rowdy (*ATD* 23)

Although Junior consistently identifies himself as a cartoonist and foregrounds the significance of drawing in his life, writing nevertheless emerges as a meaningful

⁵⁰ For instance, Forney imagined Junior drawing the “Penelope bird” by copying it out of a textbook in the school library, meditatively drawing a long time thinking about Penelope (*ATD* “Interview” 241-42).

practice as well. For instance, he reports having written love letters to Penelope (227) rather than expressing his feelings through cartoons, suggesting that written language is relevant in certain contexts. Moreover, even if Junior does not explicitly self-identify as a writer, he decidedly positions himself as a reader. After transferring to Reardan, he is inspired by Gordy to read extensively, and he shows a strong engagement with literature (“Yeah, I know, I probably read too many books,” 105).⁵¹ Given the absence of explicit motivation for the act of writing and Junior’s clearly articulated aspiration to become a cartoonist, the diary itself may be understood less as a record of written self-reflection and more as a framework that legitimates the inclusion of his drawings. The text thereby functions, at least in part, as a curated collection of his cartoons. Importantly, Junior’s stated reasons for drawing, such as escapism, emotional processing, and the representation of subjective experience, align with the motivation for writing a diary, which has been described by researchers of diary narration (cf. Kellner).

6.3.2 Multimodal Meaning-Making

Despite Junior’s passion for drawing, the majority of the novel is comprised of running text and the mode of written language is prevalent both in the running text as well as in the cartoons. Paragraphing, font, or formatting are visual affordances of written language that can emphasize certain elements and/or detach them from the running text. Spacing through paragraphs is used frequently in *ATD*, often creating a dramatic effect, e.g.

“Ten more than usual.
Ten more than normal.
Ten teeth past human.” (2)

The visual emphasis of the anaphora strongly highlights the excessive amount of teeth Junior has. Further, emphasis is achieved by employing italics or capitalizing letters throughout the running text, expressing and underscoring emotions (“my world

⁵¹ His linguistic sensitivity is further highlighted in his observations about others’ speech, such as his comments on Penelope’s and Gordy’s distinctive ways of talking (“Yeah, she talked like that,” *ATD* 111; “Can you believe there is a kid who talks like that?” 132). These remarks not only reveal Junior’s attentiveness to language but also underscore that his narrative authority derives not solely from his skills as a cartoonist but also from his capacity as a reader and acute observer of linguistic nuance.

would EXPLODE without them,” 11, original emphasis) or even emphasizing narratological tension (“And Rowdy rose into the air. And I rose with him. AND THEN I ROSE ABOVE HIM!”, 192, original emphasis). In general, there are many short paragraphs in *ATD*, adding to a more narrative than reflective stance. Especially utterances in direct speech, which are featured in an uncommonly high amount for diary novels, are separated by paragraphs. Spacing is also used at the beginning of each chapter, as the headlines are always centered vertically on a new page.

Throughout most of *ATD*, a regular computer font is used.⁵² There is little variation: for instance, “The unofficial and unwritten [...] Spokane Indian rules of fisticuffs” (61-62) employs a western font and thus appears like a passed-down document, which is noticeably not a part of the narration, as shown in Figure 23. Also, Junior includes an email from his sister in a different font (99-100, see Figure 23). A handwritten font, like in *DWK*, is only used in the cartoons in *ATD*, whereas the running text refrains from visually adapting to the diary theme. There are, however, two exceptions to this, which heavily impair the fiction of factuality, namely the inscription in Junior’s second-hand schoolbook (“THIS BOOK BELONGS TO AGNES ADAMS,” 31), and the handwritten lists of things that make Junior happy (176-78, see Figure 23). The visual presentation of the cartoons, which appear as if drawn on separate sheets of paper and taped into the diary, combined with the occasional use of handwritten font within the running text, produces a noticeable division between computer-typed and handwritten elements. This juxtaposition underscores the materiality of the typed text and draws attention to its mode of production. If the readers apply the aforementioned “principle of charity,” they might disregard the implications of the typed format, given that it seems implausible for Junior, who lives in poverty, to have consistent access to a computer. By highlighting the technological nature of the typed text, however, the novel draws attention to questions of production and thereby disrupts the fiction of factuality. This effect is reinforced by the inclusion of a handwritten letter from Junior’s sister (133, see Figure 23). It remains ambiguous whether this is a reproduction of her actual handwriting or Junior’s imitation of it. Either possibility destabilizes the

⁵² Most likely the font that is used is Times New Roman, or a very similar serif-based font.

diary's factual pretense, since he could plausibly have pasted the letter itself, copied it in his own handwriting, or merely summarized its content.

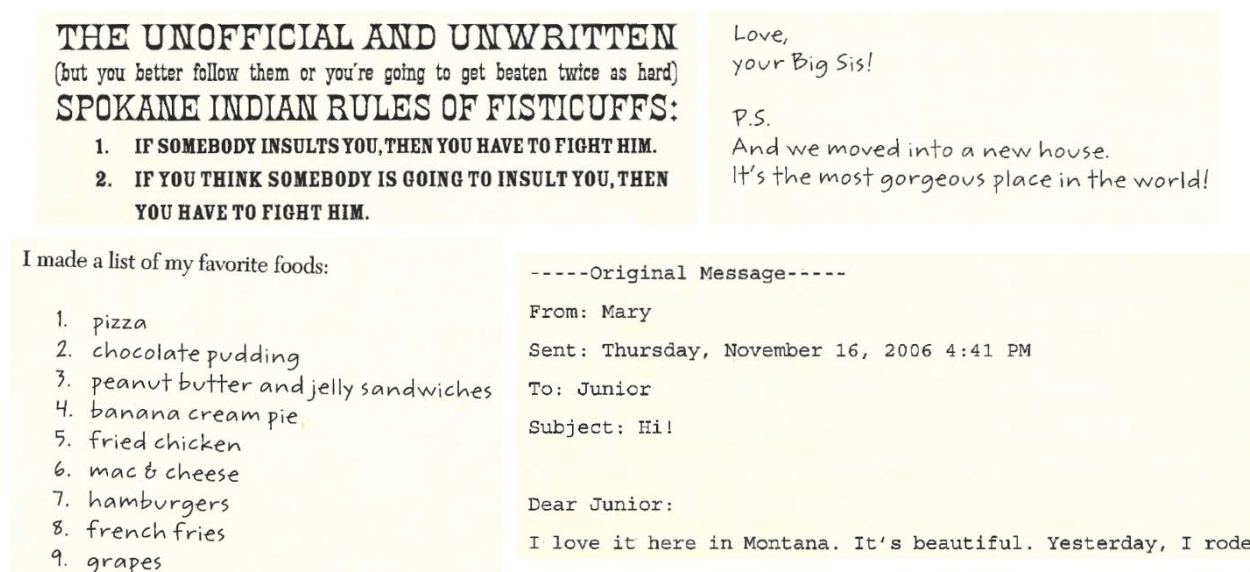


Figure 23: Different fonts in ATD (61, 133, 177, 99)

Written language in the cartoons is very prominent⁵³ and mainly serves the purpose of explanations, parodies, or annotations, for instance, when Junior draws the food at a diner he cannot afford and labels it “the pancakes of doom” (124), as shown in Figure 24. Here, the same affordances, i.e. spacing and capital letters like in the running text can be accounted for. Another important use of written language in *ATD* is to represent speech – direct speech in the running text or speech balloons in the cartoons – or sound. As far as onomatopoeia is concerned, not only do the cartoons feature typical sounds such as “SMASH!” (*ATD* 31, see Figure 25), but there are also many instances in the running text, where the onomatopoetic words are often emphasized through italics or capital letters (e.g. “*Smash! Boom! Bang!* He dented the doors [...],” 20; “We were silent, SNAP, just like that,” 137, both original emphasis). Apart from these words representing sound, Junior also imitates sound by spelling it out in the running text, e.g. “Or maybe I should say I had a st-st-st-st-stutter and a lissssssssthtthththp” (4), or when he overhears Penelope vomiting in the school bathroom: he renders the sound of it as “ARGGGHHHHHHHSGHHSLLLSKSS-

⁵³ As a matter of fact, there are only three cartoons in total that do not feature written language, two of them expressing Junior's sadness at his grandmother's funeral (*ATD* 160, 166) and one showing him and Rowdy as superheroes (102).

SHHSDKFDJSABCDEFGHIJKLMNPOQRSTUVWXYZ!” (106). The opening sequence – an auditory metaphor – onomatopoeitically imitates the stereotypical noises of retching, while the subsequent string of random letters, ending in the entire alphabet, does not represent the sound itself but rather underscores its overwhelming duration and intensity. Similarly, Junior uses written language to capture auditory experiences in other contexts, such as his description of the crowd’s reaction during a basketball game: “You want to know what it sounded like? It sounded like this: BOOOOOOOO[...]OOOOO!” (190, original emphasis). The “BOO,” extended across five lines, conveys both the volume and relentlessness of the audience’s jeers. In the cartoons as well, sound is often visualized through written cues. When Junior introduces his basketball coach, he draws a speech balloon saying “LISTEN UP” (137, original emphasis, see Figure 25). The capital letters in combination with the full stop suggest a loud and firm voice but not shouting. This is also reflected by the annotation of the speech balloon “voice like Thor” (137), invoking these qualities of spoken language. Written language in the cartoons is also used for sounds such as “cough / clank clunk clatter” (88), “grumble” (120), or “bang!” (170), as can be seen in Figure 25.



Figure 24: Labelling in the cartoons (ATD 124)



Figure 25: Representation of sound in ATD (137, 31, 88)

Additionally, laughter as a sound plays an important role throughout the book, both in the running text as well as in the cartoons. For example, Junior describes Rowdy’s

laughter, which he likes a lot, and compares it to an “avalanche of ha-ha and ho-ho and hee-hee” (*ATD* 23). Figure 26 shows Junior on the car ride home after hearing about his sister’s death. Junior laughs as a coping mechanism, and the sound literally fills his side of the car with “ha” written in various fonts and sizes around him (204).⁵⁴ This prominent role of laughter in the novel resonates with contemporary Native American literary traditions, where laughter, or more generally humor, often functions as a subversive strategy and a form of resistance and survivance (cf. Kertzer; Bouzonviller). The novel includes both, an intercultural dimension of subversive laughter, but also a form of “intracultural” laughter, where Nativeness is negotiated within Native communities (Gruber 3).



Figure 26: Laughter in *ATD* (204)

Although Junior sees himself as a cartoonist, rather than a writer, and aspires to earn recognition in this field, he is very proficient linguistically. Despite the casual and conversational tone of the narration (“Well, anyway, [...],” *ATD* 182, as if losing his trail of thought; “Ha-ha, just kidding,” 190), he employs many literary devices skillfully, like personification (“So destiny and me lean against the wall and wait,” 106) or similes (“My clothes and hair were soaked with her tears. It was, like, my mother had given me a grief shower, you know? Like she’d baptized me with her pain,” 208), to only name a few examples. He also uses irony often, especially when talking about himself or his situation (“Yeah, I was so *articulate*,” 84, original emphasis). The narrative style

⁵⁴ This focus on sound in the novel foregrounds the tension between spoken and written language. On his first day at Reardan, his classmates are unsettled by his speech. Junior explains, “I had that stutter and lisp, but I also had that singsong reservation accent that made everything I said sound like a bad poem” (*ATD* 61). While such phonetic qualities cannot be conveyed through writing, their explicit mention highlights the limitations of the written medium and reminds readers that the many direct quotations in *ATD* inevitably lack certain oral dimensions of Junior’s voice.

of the running text is characterized by immediacy and spontaneity. Junior's remark, "Wow, he didn't kick my ass" (72), for instance, conveys an unfiltered reaction that blurs the line between the experiencing and narrating I, a quality atypical of conventional diary novels. This sense of immediacy and conversationality is reinforced by frequent paragraphing, often including isolated single-sentence paragraphs, as well as the frequent use of colloquial expressions ("jeez," 2; "you know," "Yep," 3), which give the text a distinctly oral quality. With regard to orality in Native American literature, the incorporation of oral traditions has been examined, for instance, by Teuton, who critically discusses an assumed orality-literacy dichotomy. While this debate is highly relevant, a more detailed analysis would exceed the scope of this thesis. Likewise, a comprehensive analysis of oral traditions in *ATD* cannot be undertaken here. Generally, Alexie is frequently described as drawing on oral storytelling techniques, particularly in his short prose (cf. Ibarrola Armendariz). What can be said for *ATD*, is that acts of speaking, i.e. dialogue, play a central role in the novel, and storytelling is highlighted over diary-typical reflection. In this way, Junior's verbal narration seems very deliberate, particularly in his strategic use of pacing to create suspense and emotional impact. The deaths of close friends and family members, for example, are often revealed abruptly and embedded within otherwise mundane accounts, producing striking juxtapositions. These sudden shifts in narrative pace mirror the unexpectedness of these losses, confronting the reader with their shocking immediacy.

Although the verbal mode quantitatively outweighs the visual mode, the cartoons in *ATD* are a central aspect of the narration. There are only sixty cartoons distributed across 230 pages, which amounts to roughly one illustration on every fourth page. The cartoons themselves are heterogeneous in both size and style, ranging from full-page illustrations to small insertions, and varying in their degree of cartoonicity as well, from more detailed, finely penned drawings to simplified, bold-lined caricatures. A further noteworthy aspect is their material presentation. Many cartoons appear to be drawn on paper different from that of the diary itself, frequently depicted on lined sheets with punch holes or on blank, crumpled pages with torn edges. These drawings are typically shown with two or three strips of tape affixing them at opposite corners, thereby creating the impression that Junior produced them elsewhere (e.g. in a

separate notebook), tore them out, and subsequently pasted them into his diary. Significantly, however, the narrative never explicitly comments on this practice, leaving the integration of text and image to visual implication alone. By contrast, the full-page cartoons depicting characters or the comic-strip style sequences lack the illusion of being physically inserted.

Junior portrays both objects and people with varying degrees of iconicity, depending on the drawing process, as described above. For example, as Figure 27 illustrates, his mother is featured in a more detailed cartoon, highlighting certain physical features about her (ATD 12), as well as a comic strip, where the iconicity – i.e. abstraction – is higher and her nose is omitted, for example (174). The level of iconicity can also vary within one cartoon, as can be seen in Figure 28. For instance, the drawing on page 63 shows Junior, displaying all the iconic signs of fear (e.g. wide eyes, shivering indicated by lines next to his body), surrounded by giant uniformly shaped bodies with only two dots for eyes and angry (i.e. v-shaped) eyebrows, calling out racist slurs. This creates and emphasizes an intimidating effect.



Figure 27: Varying levels of iconicity across cartoons (ATD 12, 174)

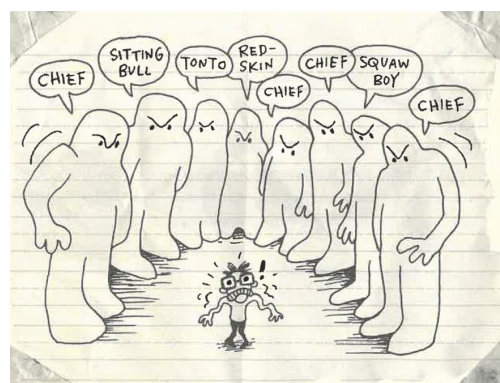


Figure 28: Varying levels of iconicity within one cartoon (ATD 63)

The drawings employ a typical visual vocabulary of cartoons, as illustrated in Figure 30, such as different lines symbolizing motion (e.g. a throwing motion, 31), or zig-zag lines symbolizing anger (31). Drops symbolizing sweat beads and thus an exhausting activity (e.g. 88) or embarrassing situations (e.g. 128) form contrasting intersemiotic relations across the novel (cf. Packard et al. 61), and bubbles emerging from drunk people's heads, indicating their inebriation (e.g. ATD 170 and 213), form parallel relations (see Figure 29). Figure 29 and Figure 30 show even more visual metaphors, such

as exaggerated facial expressions (e.g. 142), zig-zag shapes symbolizing explosions or gun shots (170), sharp teeth symbolizing anger (31), or xs for eyes symbolizing death (213), and further illustrate emotions and states of being, showing Junior's subjective perspective. Further, visual metaphors are also vital for the expression of Junior's subjectivity, as seen in Figure 30. For example, Junior does not express verbally what he thinks of Penelope's racist father, but portrays him as a slimy and disgusting worm (109). At his grandmother's funeral, he explains how "laughter and tears are pretty much the same thing" (166) and offers a creative visual representation of this complex feeling. Another relevant visual metaphor is the depiction of the "Australian tufted Arnelope" (ATD 112), a bird Junior draws in great detail to visualize Penelope's and his desire to leave their current situation.



Figure 29: Contrasting and parallel intersemiotic relations in ATD (88 and 128; 170 and 213)



Figure 30: Visual metaphors in ATD (31, 142, 170, 112, 166, 109)

Regarding the reading path, the predominance of verbal narration means that many pages contain no cartoons, so the text largely guides the reader. When cartoons do appear, they usually span the page width and are embedded without disrupting the flow, rarely interrupting a sentence. Since most images do not advance the narrative, readers quickly learn that their order of reception, whether text first or cartoon first, is flexible, which also accounts for full-page portraits or comic strips. Since the cartoons do not further the narration, they might be seen as interruptions thereof and sometimes they ‘interrupt’ two sentences like for example the end of a page would (e.g. *ATD* 90-91, see Figure 31). Overall, the reading path is relatively open, though occasional multimodal ties create moments of structured guidance. There are different kinds of multimodal ties. Some cartoons are connected to the running text with a colon and explicit deixis including demonstrative pronouns (“And this is what I saw when I stared at Penelope;,” *ATD* 113). Ties can also be established exclusively through the use of a colon, e.g. the six-panel comic strip titled “How to Get the Last Sip of Wine from the Bottom of the Bottle” (170), which shows alternatives to Eugene’s death, is introduced without deictic elements, but only through a colon (“The police think Eugene and Bobby fought over the last drink in a bottle of wine;,” 169). Weaker ties leave greater interpretive space and may disrupt coherence. For instance, the cartoon of a bruised bottom annotated “all ass-kicking bruises look like Texas...” (45) is only loosely connected to the preceding line (“I felt like life had just knocked me on my ass,” 44) and functions more as a humorous aside than as a direct extension of the narration.

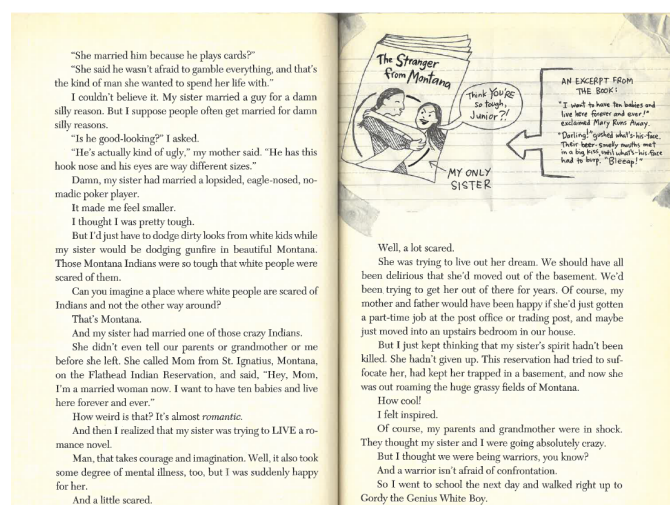


Figure 31: Example of page layout (*ATD* 90-91)



Figure 32: Direct speech in the cartoons (ATD 96)

While direct speech is a vital part of the running text, the presence of speech balloons in the cartoons is rare and usually does not present actual conversations, but rather exaggerated comments. There is only one conversation that is represented as a cartoon with split speech balloons: Figure 32 shows how Junior and Gordy talk about their love of books (ATD 96), which is the only conversation that does not entail past tense tags (e.g. “he said”).

In addition to character portraits, *ATD* contains five full-page comic strips composed of six panels each, with the first panel presenting the title. Like the other cartoons, these strips lack a clear narrative function, either within the overall plot or as self-contained sequences. Instead, they present non-consecutive scenarios that highlight alternatives (“How to Get the Last Sip of Wine from the Bottom of the Bottle,” 170), illustrate possible actions (“How to Pretend You’re Not Poor,” 120, see Figure 33), or emphasize repetition (“Why I Did Actually Miss a Lot of School,” 174).



Figure 33: Six-panel comic strips in *ATD* (120)

Regarding the functions of cartoons, expressing humor is an important feature of Junior’s diary and especially his drawings, as shown in Figure 34. On the one hand, he uses humor as a coping mechanism by mocking the events that happened to reduce their severity. After his grandmother and Eugene die, Junior feels taunted by God and Jesus and uses his cartooning to ridicule them in return, by portraying Jesus’ disciples laughing at him because he “farteth and burpeth in harmony MIRACULOUS!! John 11:35 am” (ATD 171). On the other hand, Junior in general uses his cartoons “to make fun of the world [and] people” (95) and express his humorous side. For example, he

draws “The Eimouttaheer Virus,” (138) after more and more students quit the basketball tryouts, as if quitting were contagious. He also makes fun of himself, as evident by his self-portrait at the beginning of the book titled “Me in all my Glory” (5, see Figure 35), showing him with exaggerated physical features in a non-flattering way, mocking his speech impediments in a speech balloon and his epilepsy through a shaking body. Although Junior’s humor often relies on the presence of words, he mainly uses his cartoons to mock something or someone (e.g. his caricature of his report card, that includes mocking puns on subjects, 216). He often makes comparisons (see Figure 36) that are exerted both through drawing and writing, e.g. a Native American dancer and a chicken (19) or his teacher and a volcano (85). He caricatures, for example of the romance novels his sister reads (38, see Figure 37). He also uses his cartoons to mock people, for example Ted, the white man appearing at his grandmother’s funeral, claiming to “feel Indian in my bones” (162, see Figure 35). While Ted’s appearance looks rather normal, the annotations pointing to different clothing items clearly portray the man in a ludicrous manner: “‘sacred leather scrotum sheath’ purchased from Navajo shaman for \$1000 (actually a Kmart Naugahyde baby booty purchased by Navajo conman for \$3.99)” (162).



Figure 34: Visual mocking in ATD (171, 138, 216)

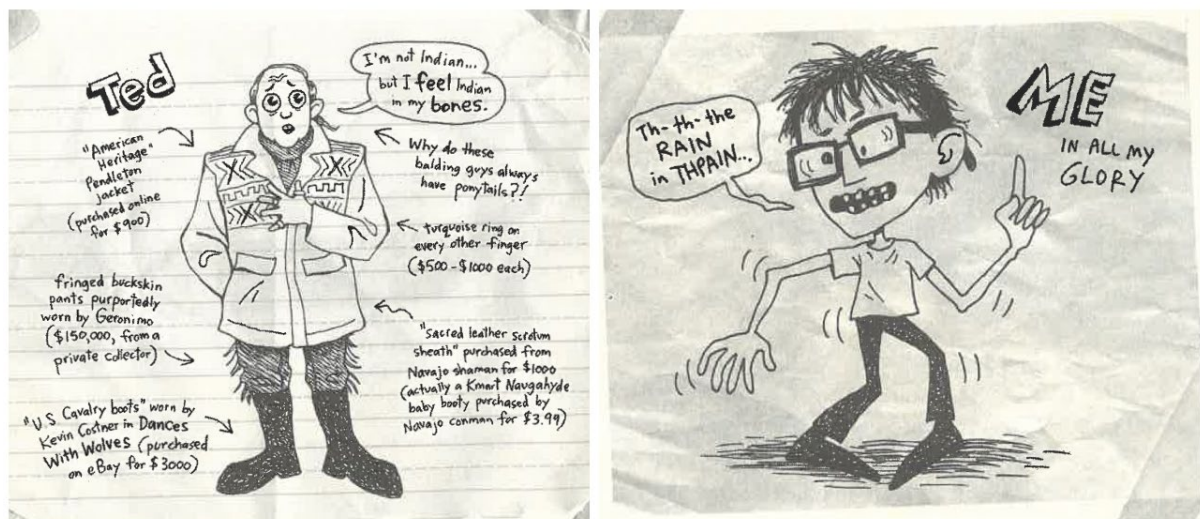


Figure 35: Mocking people (ATD 162, 5)

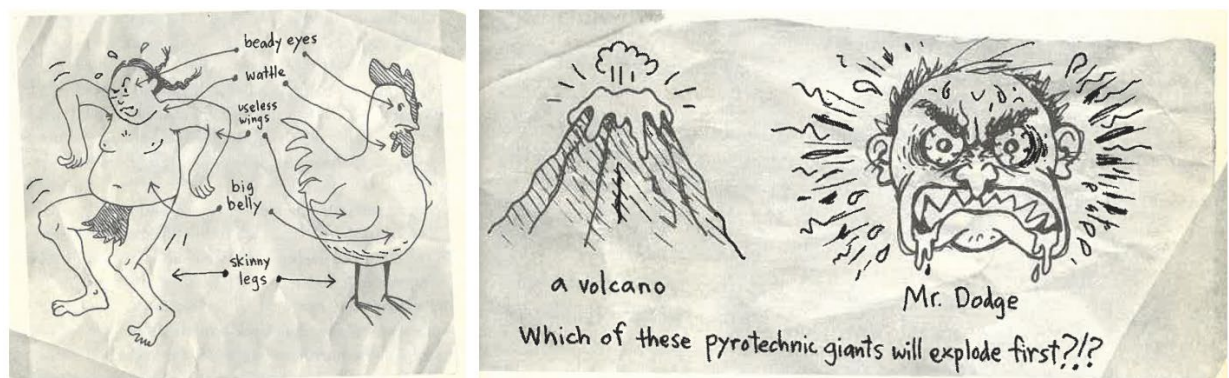


Figure 36: Visual comparisons (ATD 19, 85)

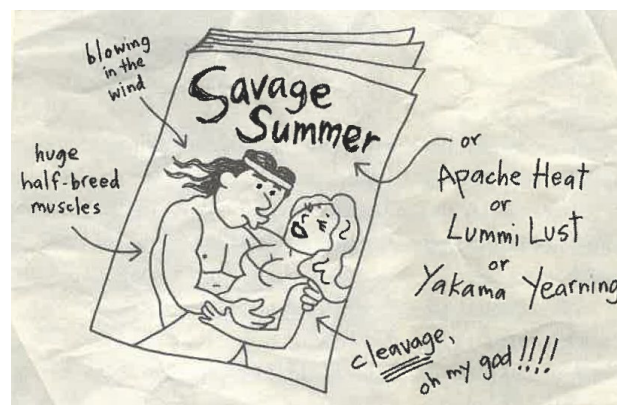


Figure 37: Caricature (ATD 38)

Apart from mocking people in his cartoons, Junior also claims he draws to honor them (ATD 95). The portraits of his loved ones are used to introduce and characterize his friends and family by wittily commenting on different physical features or items and clothes they wear, for example his grandmother in Figure 38 (ATD 69). There are annotated portraits of his parents, grandmother, sister, his teachers (Mr. P, Coach) and Ted.

In terms of honoring the people he loves, there are also rather detailed, more time-consuming portraits with fewer and non-ironic annotations of his friends, namely Rowdy, Gordy, Eugene (see Figure 38), and Penelope. The iconic portraits usually span across an entire page, apart from those of Mr. P and Ted, and seem to not have been inserted, whereas the penciled, less iconic, ones, also have been drawn on a paper different from the novel background. The difference is that these portraits of beloved ones do not have a torn outline, but seem to have been cut out, which shows the appreciation Junior has for these people, which is illustrated in Figure 38, showing Eugene's and Mr. P's portrait. The fact that Mr. P and Ted are featured in maximum half-page cartoons and have a teared outline, shows who Junior really appreciates and makes the distinction between his two teachers, Mr. P and Coach, clear. Thus, Junior's subjective feelings are transported through multiple modes.



Figure 38: Portraits in ATD (69, 71, 29)

Another important function of the cartoons in the novel, apart from mocking or honoring people, is visualizing emotions, which is illustrated by Figure 39. For instance, a cartoon of Grandmother Spirit's funeral shows Junior's face with eyes closed, crying, and surrounded by three Native Americans forming a circle around him (ATD 160).

Apart from sadness, this cartoon also transports a sense of community, portraying a layered merging of emotions. Given their immediacy, the cartoons are able to intensify emotions and draw the reader in. A prominent example of this is when Rowdy yells at Junior for attending school outside of the reservation (53). The frontal view of Rowdy's outraged face makes the readers feel like the addressee of his anger. While the running text moves along the narration and describes emotions to a degree, the immediacy and virulence of the emotion portrayed in a cartoon is higher and draws the reader into the moment where a friendship (temporarily) ends.



Figure 39: Visualizing emotions in *ATD* (160, 53)

Another significant function of the cartoons, frequently discussed in the secondary literature, is their role in identifying and reflecting cultural stereotypes (cf. Schlaffer; Kertzer). Rather than advancing the plot, the images are often symbolic, serving as a creative outlet for Junior's reflections on his struggles. A prominent example is the iconic "part-time Indian" cartoon (*ATD* 57, see Figure 40), where Junior's body is divided into a "white" and an "Indian" half. According to Suhr-Sytsma, this image has been misread by critics as depicting a split identity, when in fact it rather critiques white privilege. After all, the white side lists economic opportunities and social advantages, such as "hope" or "positive role models" that remain unattainable for Junior, underscoring the structural inequality that shapes his life (Suhr-Sytsma 36).⁵⁵ Similarly, the cartoon in which Junior depicts himself as a stereotypical warrior, complete

⁵⁵ Other scholars have pointed to the legacy of the Carlisle Indian Industrial school ideology of "kill the Indian to save the man," which aimed at erasing Indigenous identity through forced assimilation. As Kertzer argues, this cartoon visually encapsulates these two opposing sides, a rejected "before" and a desired "after" (61). This historical agenda is also directly addressed in *ATD*, where Mr. P admits that, in his early years of teaching on the reservation, he had been expected to implement precisely such assimilationist policies (*ATD* 35). For a deeper analysis of how this type of 'education' is represented in *ATD* see Katanski.

with feathered headdress and loincloth (ATD 142, see Figure 30), demonstrates his awareness of the reductive expectations imposed on him at his predominantly white school. By drawing himself this way, Junior both engages with these stereotypes and simultaneously attempts to reclaim the warrior identity, depicting himself shouting with confidence and energy (cf. Szeghi; Schlaffer). This cartoon suggests that the stereotypical constructions of Indianness override his individual identity, influencing how he understands himself and how he imagines others perceive him (cf. Ashley). In this way, multimodality is crucial for exposing and critiquing dominant stereotypes far more immediately than a ‘verbal only’-narration.

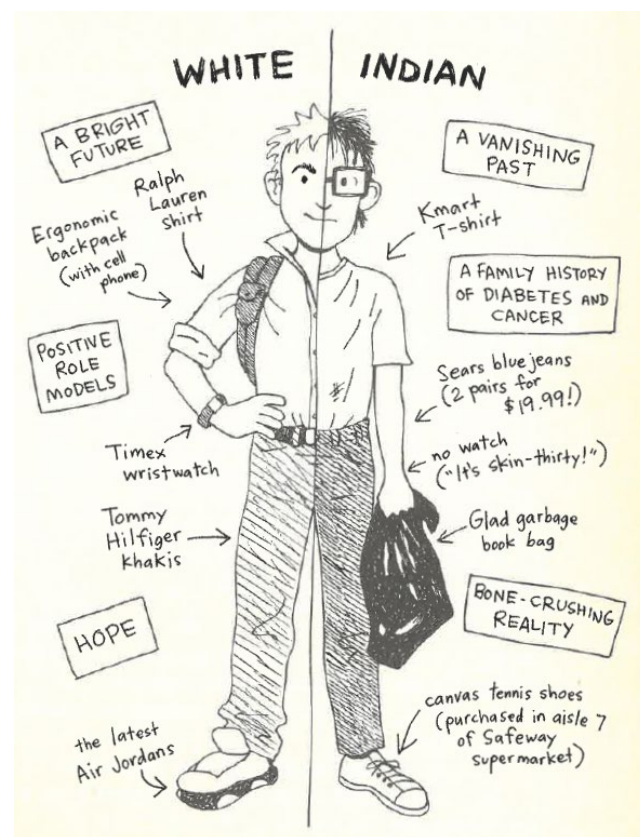


Figure 40: Reflecting cultural stereotypes (ATD 57)

It seems evident that in *ATD* the primary relation between text and image corresponds to what McCloud terms a “word-specific” (153) combination. This means that the verbal mode, both in the running text as well as in the written elements embedded in the cartoons, largely guides the meaning-making of a narrative. It suggests that the visuals then merely serve an illustrative purpose, as is also suggested by McCloud. While this is certainly true for some cartoons in *ATD*, for others it seems so only at first

glance. Some cartoons contribute significantly to the characterization of Junior, his drawing process, his self-perception, as well as his feelings and attitudes. Moreover, they frequently introduce moments of reflection, whose nuance goes beyond the information provided by the verbal narration alone.

6.3.3 Constructing Narrative Stance

The verbal narration in *ATD* conventionally takes the form of an autodiegetic diary voice, with Junior recounting his own experiences and impressions. At the same time, the text allows for rare diegetic shifts, most notably through inserted letters and an email from Junior's sister, which provide first-person accounts of her life after marriage and her move out of state.

With regard to focalization, the verbal narration in *ATD* follows the conventions of diary fiction, as Junior consistently functions as the primary focalizer. This also extends to the cartoons, which largely reflect his subjective perspective, often providing ironic or comedic commentary rather than representing events in the story world directly. They visualize Junior's thoughts, perceptions, and metaphors, such as his depiction of his brain as a patchwork quilt annotated as "the brain of many colors" (*ATD* 87), and thus offer insight into his inner life. The cartoons frequently convey how Junior perceives others, for instance in the annotated portrait of his sister, where he notes her "acne scars that somehow make her look tough & pretty at the same time" (27). At times, the images represent an optical first-person perspective, such as the cartoon of Rowdy angrily confronting him about transferring schools (53, see Figure 39). Through these subjective representations, the cartoons reinforce the internal focalization of the narrative, occasionally creating explicit ties, as cited above.

In addition to internally focalized cartoons, *ATD* also features instances of external focalization. These can be categorized into three subtypes: (1) reflective cartoons that are subjective but not part of the storyworld, (2) cartoons that are subjective but situated within the storyworld, and (3) a rare case of intersubjectively valid and narrative external focalization. The first and second categories occur throughout the novel, while there is only one instance of the third type.

The first subtype includes reflective, non-diegetic depictions of Junior's emotions. For instance, at his grandmother's funeral, he is shown encircled by a crowd of Native Americans (*ATD* 160, see Figure 39), a visual metaphor that captures his feelings of grief and community rather than portraying an actual scene. Similarly, the cartoon in which he is surrounded and harassed by students at Reardan (63, see Figure 28) reflects a real event but exaggerates it iconically to emphasize emotional intensity rather than giving an objective account of the incident.

The second subtype comprises cartoons that remain grounded in the storyworld yet still foreground Junior's subjective perception. One example is a drawing of himself laughing hysterically when his father arrives after his sister's death (204, see Figure 26). Although the running text conveys his feeling of "going absolutely insane with laughter" (206), the cartoon exaggerates this state, by showing Junior with a spinning head and a maniacal facial expression, surrounded by "HA-HAs" in various sizes filling up the passenger side of the car. This offers a more visceral representation of his perception. Likewise, the cartoon of Junior throwing a book at Mr. P (31, see Figure 30) dramatizes the event in a way that is not literally possible, exaggerating both Junior's anger and Mr. P's fall to underline the emotional stakes rather than the factual sequence.

The final subtype, an intersubjectively valid external focalization, means that a scene is depicted objectively, showing Junior from the outside. Such an instance occurs only once, in the conversation with Gordy about books (*ATD* 96, see Figure 32). Unlike most cartoons, this one is not merely reiterative of the verbal narration but depicts their dialogue directly in speech balloons, thus posing the only interdependent combination (cf. McCloud 155). Although Junior announces the scene with "I was shocked" (*ATD* 96) the cartoon does more than show his feelings, as it captures the interaction itself without repeating the running text, thereby achieving a narrative quality beyond subjective commentary.

Although there are shifts in focalization, particularly through multimodality, with the running text providing autodiegetic narration while the cartoons often present an externalized view of Junior, the overall text, in both modes, remains highly subjective. Junior engages in extensive verbal reflection, which the cartoons reinforce by

visualizing his self-perception and sense of irony. His drawings frequently emphasize his own awkwardness and vulnerabilities in an ironic way. The very first cartoon portrays him “in all his glory” (ATD 5, see Figure 35), exaggerating his oversized head, trembling limbs, and stuttering speech, while other cartoons highlight moments of insecurity, such as attending the Winter Formal in his father’s outdated suit (121, see Figure 41). Through these depictions, the reader not only gains insight into how Junior views others but also into how he sees himself, often in a self-deprecating light. This particular form of humor is especially prominent in Native American literature and is frequently described as one of its most characteristic features (e.g. Gruber 200). As Jannetta argues, self-deprecation often functions as a “pre-emptive strike” (122), anticipating ridicule in order to deflect it and thereby mitigate potential humiliation. The cartoons further enrich his characterization by providing visual details absent from the verbal narration, such as his glasses or hairstyle. Overall, Junior’s style is marked by irony, as he consistently uses both words and images to satirize not only those around him but, above all, himself and his situation. Satire is likewise a central strategy in Native American literature, combining sharp social criticism with an opportunity for what Gruber calls “healing laughter” (211). It enables Native authors to challenge stereotypical notions of Indigeneity and expose the biased assumptions underlying dominant representations (Gruber 61–62).

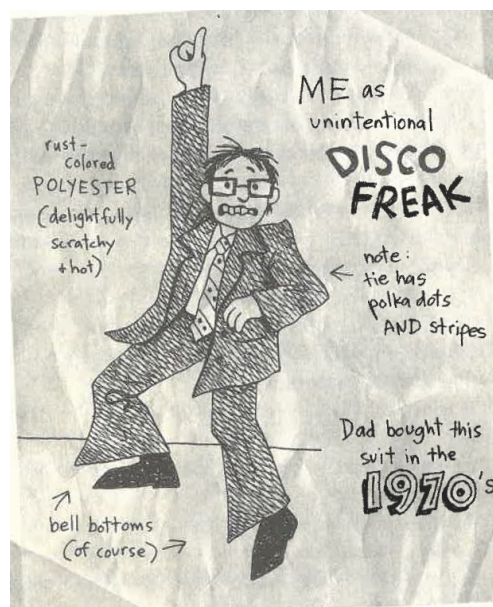


Figure 41: Self-deprecation (ATD 121)

It becomes evident that Junior identifies not only as a cartoonist but as a reader. This dual self-conception is reflected in the high degree of subjectivity throughout the novel, conveyed both verbally and visually through direct and indirect characterization. Junior describes himself as “a freak because I loved books” (ATD 98), and after transferring to Reardan, Gordy’s encouragement further deepens his engagement with literature. His reading extends beyond enjoyment to shaping his narrative style, as seen in his chapter titles, which sometimes refer to canonical literature, for instance, “Because Russian Guys Are Not Always Geniuses” (199), which links his own experiences to a discussion with Gordy about a Tolstoy novel. After he starts attending Reardan high school, his cartoons further highlight this literary orientation by visually representing major life events as book covers, as shown in Figure 42. When his sister elopes with a Flathead Indian, he depicts it as a novel called *The Stranger from Montana* (91). Also, his budding romance with Penelope becomes *The Other Side of the Fence* (108), and after Mary’s tragic death, he creates the cover for *Burning Love* (213), where he features his sister on the cover, engulfed in flames, directly addressing him with “Thanks a lot, Junior.” These examples illustrate how Junior’s identity as a reader intersects with his visual creativity, merging his literary imagination with his cartoonist practice.



Figure 42: Book covers (ATD 91, 108, 213)

Another central theme in Junior's subjective self-perception is the recurring sense of division, echoing the novel's title of "part-time Indian." This feeling of being split is articulated both verbally and visually across the text. One striking example appears before the rematch against Wellpinit, where Junior depicts himself simultaneously as an angel cheered by Reardan and a devil booed by Wellpinit, captioned with the question, "Who am I??" (ATD 182, see Figure 43). While the verbal narration already conveys his sense of fractured identity ("I was half Indian in one place and half white in the other. It was like being Indian was my job, but it was only a part-time job," 118), the cartoon visualizes this feeling of division, intensifying the representation of his conflicted self-conception. This interplay of verbal and visual modes highlights how multimodality enables the novel to dramatize Junior's inner conflict.



Figure 43: Split sense of self (ATD 182)

Furthermore, multimodality plays a crucial role in reflecting and interrogating stereotypes, both by exposing them and by allowing Junior to articulate his own entanglement with them. As Garić observes, Junior himself acknowledges that he internalizes stereotypes, for instance when he objectifies Penelope, desiring to be the "chocolate topping" to her "vanilla dessert cake" (ATD 114), a view his friends Rowdy and Gordy immediately condemn as racist (Garić 191–92). Similarly, when Junior depicts himself in a loincloth and feathered headdress (ATD 142, see Figure 29), the cartoon does not represent his actual appearance but rather his own internalized image of the stereotypical Indian. As Ashley argues, such moments demonstrate how Junior's sense of self is overlaid by culturally imposed stereotypes. His Indian-ness is perceived less as an individual identity than as a cultural construction shaped by external expectations (Ashley 29).

6.3.4 Findings

At first glance, the title suggests that *The Absolutely True Diary of a Part-Time Indian* should be read as a diary novel. However, the text departs in significant ways from the proto-typical conventions of this genre. As genre plays a crucial role in multimodal analysis, this deviation is particularly noteworthy, as *ATD* reshapes some of these norms, primarily through its verbal mode and overall narrative tone. Unlike most diary novels, which foreground the temporal gap between the experiencing and narrating I, *ATD* shows a highly conversational and dynamic style that resembles a conventional first-person narrative, blurring the distinction between Junior as a diarist and as the protagonist of his own story. Long stretches of dialogue, often spanning entire pages, create the impression of immediacy, almost conflating discourse and story time. While the tone of the narration is reflective, direct speech and onomatopoeia in the running text (“GRRR,” “HA-HA”) further strengthen the sense of oral immediacy. This way, the verbal narration gains a *showing* quality, deviating from the conventional *telling* quality of diary novels in general. The narrative, thus, seems more engaging for the reader due to its liveliness and dynamic.⁵⁶ Yet, the multimodal interplay between text and image reintroduces diary-like qualities. While the running text leans toward narrative fiction, the cartoons foreground the reflective and subjective dimension of Junior’s storytelling, highlighting him as the narrating I. They provide often humorous, ironic, or self-deprecating, commentary that draws attention back to the narrating I and the act of self-representation. In this way, the images restore some of the introspective quality typical of diary novels, even as the verbal mode serves a more narrative purpose. While Junior does reflect verbally, this is often done from a point of the events as they take place in the narrative, where the experiencing I already reflects on his situation, which is reinforced by phrases like “Ain’t that weird?” (*ATD* 205).

Narratologically, the novel also employs strategies that undercut the diary illusion. By withholding or delaying information, as in the funeral scene, in which both Ted and the reader are surprised by the revelation about Grandmother Spirit, Junior manipulates narrative perspective in a way that feels crafted rather than diaristic. The diarist

⁵⁶ This reveals that *ATD* is in fact a YA novel, which is meant to be entertaining and appealing to a broad readership, so it can be sold to a large audience, rather than a personal record of events and means of reflection.

deliberately leaves the reader with just as little knowledge as Ted to create narrative tension. This way, the reader is as surprised as Ted when Junior's mother reveals that Grandmother Spirit, albeit being „powwow-famous“ (*ATD* 156), actually never danced in any of them. These instances of atypical diary narration are reinforced by the fact that there is no clear addressee and explicit intention for writing the diary, elements that might otherwise reduce such moments of doubt.

Taken together, *ATD* is best described not as a diary novel in the strict sense but as an autofictional first-person narrative enriched by multimodality. The reflective cartoons, as well as the blending of narrating and experiencing perspectives reveal that, ultimately, the text is less about keeping a diary and more about telling a story that entertains, reflects, and negotiates identity through the interplay of words and images.

6.4 Comparative Discussion

In both *DWK* and *ATD*, the running text is characterized by an autodiegetic voice, as the diarist functions simultaneously as narrator and protagonist. When the cartoons include verbal commentary, as is particularly frequent in *ATD*, they, too, adopt an autodiegetic voice, thereby reinforcing the diarist's subjectivity. What stands out in both texts, however, is the extensive use of direct speech, which is atypical for the diary novel as a genre. Yet, this cannot be attributed solely to the multimodal dimension of cartoons. While in *DWK*, direct speech occurs predominantly within the cartoons, leaving the running text more conventional, *ATD* integrates direct speech extensively into the running text itself, which is a deviation from diary-novel norms. Furthermore, both texts employ diegetic shifts through inserted communications, such as letters or emails from other characters, which introduce moments of heterodiegetic narration. This demonstrates how multimodality not only provides the diarist with multiple expressive channels, but also enables the inclusion of different voices. In this sense, multimodality facilitates the voicing of other characters, momentarily reducing the diarist's authority, but still, the overall tone of the novel (highly conversational in the case of *ATD*) remains decisive in shaping how strongly this effect is felt. On one level, such moments weaken the hegemonic dominance of the diarist's narrative voice, a

feature particularly relevant in *DWK*, where the difference between the single-voiced text and the multi-voiced cartoons is crucial for exposing Greg's biased and unreliable self-presentation. Thus, the presence of additional voices in both texts diminishes the extent of unreliability typically associated with diary narration, since the story is not mediated exclusively through the diarist's perspective. Although this effect could also be achieved in a monomodal setting, the blending of modes can ultimately challenge the genre conventions of diary narration, which traditionally emphasize a subjective and reflective mode of telling rather than a dramatized and dialogic rendering of events.

A constitutive feature of diary narration is its reliance on internal focalization, as the diarist's perspective is the default point of view and his:her subjective experiences and reflections are the main topics in this genre. The running text in both novels largely adheres to this principle, thereby conforming to the genre convention. However, the cartoons often deviate from this fixed standard by introducing external focalization and occasionally offering intersubjectively valid viewpoints. These shifts are particularly significant because they create narrative variation and mitigate the predominance of the diarist's perspective. In *DWK*, for example, this contributes to exposing Greg's unreliability as narrator. By consistently presenting him from the outside, without the privileged access to his inner thoughts that is reserved for his verbal self-narration, the cartoons put him on the same level as other characters, who otherwise remain subject to his biased account. This external perspective thus produces a form of equality among characters within a narration otherwise dominated by the diarist's selective control. The resulting discrepancy between what Greg says about himself and what the cartoons show subtly uncovers his lack of empathy, and his unreliability as a friend. In this respect, the shift in focalization becomes crucial for the reader to gain a fuller and more reliable understanding of the diarist. At the same time, this divergence generates much of the novel's humor and is staged in a way that enhances its entertainment value, which has undoubtedly contributed to the series' widespread success.

Generally, by including iconic depictions in the novels, a broader reader identification is fostered. The cartoony style simplifies representation and uses a visual vocabulary

that is particularly effective for conveying emotions, providing an immediate and accessible overview of a situation. It is noteworthy that, in this way, multimodal texts not only broaden the range of focalization but also enhance their marketability. The iconic, easily recognizable style increases engagement and ensures identification across diverse readerships, thereby serving, perhaps, the commercial aim of appealing to as wide an audience as possible.

Focalization in cartoon-diaries, however, does not necessarily have to shift under the influence of multimodality. In many instances the internal focalization typical of diary narration continues, only this time translated into a visual register. Such moments usually do not provide intersubjectively valid perspectives but instead foreground inner states (especially in *ATD*), attitudes, and reflections through the use of visual metaphors and iconic imagery. The cartoons thus extend the diarist's subjective perspective into another semiotic mode, enabling commentary on particular situations, people, or objects. These depictions can be closely tied to the narrative situation described in the running text and occasionally show an outside view of the diarist, but they can also be more imaginative or symbolic, representing internal states that might otherwise remain unarticulated. In contrast to externally focalized cartoons, which often highlight the diarist's self-presentation from the outside, internally focalized cartoons tend to emphasize the diarist's perception of their environments and the way they see others. Both types, however, serve to reinforce core features of diary narration, namely subjectivity and self-reflection. Multimodality, especially in *ATD*, does not disrupt the genre convention but rather strengthens it. By visually expressing internal focalization, even if the diarist gives an external, but subjectively shaped, perspective of himself, the texts both underscore the diarist's dominance as focalizer and intensify the reflective mode that is constitutive of diary novels.

Internally focalized first-person visualizations appear only as rare exceptions in these texts and it seems that they are not typical for cartoon-diaries. This is striking because, at the level of verbal narration, internal focalization in the form of autodiegetic narration constitutes the default mode and is essential for establishing the diary novel's genre identity. In other words, the running text consistently conveys the diarist's thoughts, feelings, and reflections from within, making internal focalization a

constitutive diaristic feature of the written mode. By contrast, the visual mode only occasionally adopts this autodiegetic perspective, more often presenting the diarist from an external viewpoint or offering generalized, intersubjectively valid depictions. This distribution underscores the fact that verbal narration is more apt at representing inner states, as words are particularly suited to express introspection and subjective reflection. In contrast, visual narration tends to emphasize exteriorized perspectives, as images naturally gravitate toward outward visualization and iconic condensation. In multimodal diary novels, then, the written and the visual complement each other by privileging different focalizing functions. Verbal narration anchors the subjectivity required by the genre, while visual narration expands the range of perspectives by foregrounding externalization and iconicity.

Thus, subjectivity in multimodal diary novels proves to be far more dynamic than the conventions of the genre would suggest, as the multimodal context enables multiple configurations of subjectivity. Despite external focalization, as is evident particularly in *DWK*, visuals can show a lived present, which, despite not conveying inner states, carry subjective weight through their immediacy and iconic condensation. Evidently, multimodality also allows for the representation of explicitly subjective perspectives, such as Junior's cartoons in *ATD*, which often serve as reflective or ironic commentary on narrated events. The contrast between the two novels highlights how different modes can show subjectivity: While *DWK* features a highly subjective running text complemented by more externally focalized visuals, *ATD*, with its high amount of dialogue in the running text and temporal distance between events and narration, relies more heavily on the cartoons to provide subjectivity. These differences demonstrate how, in multimodal texts, subjectivity does not rely on one mode alone but is instead negotiated between them, with each mode fulfilling distinct yet complementary functions.

In conventional diary novels, the narratorial stance is mainly reflective, as the diarist typically writes in response to a personal crisis, using the diary to vent emotions, contemplate experiences, and process feelings. This tendency is also evident in the two works under discussion. Both Greg and Junior are outsiders and lack friends and popularity, while facing crises of different magnitudes: Greg struggles with the challenges

of entering middle school, while Junior is confronted with far more existential issues concerning his family, his future, and his identity. In cartoon-diaries, however, reflection and emotional expression do not necessarily happen exclusively, or even primarily, through the verbal mode. Instead, reflection can be distributed across modes. *ATD* combines verbal narration with visual commentary, allowing the cartoons to extend, supplement, or even replace reflection, for example when Junior's drawings offer ironic or critical insights into his situation. By contrast, *DWK* seems less reflective, although this is mainly because of Greg's apparent inability to do so, providing rather superficial accounts of his feelings in the running text, while the cartoons remain predominantly narrative.

The overall tone of both works is shaped by humor, as well as sarcasm or irony, which seems to be a notable feature of cartoon-diaries. While humor is not inherently dependent on multimodality, the combination of word and image provides an especially fertile ground for it, and in these texts the interplay of modes amplifies and reinforces the humorous stance.

Furthermore, the degree and form of reflection appear closely tied to how the processes of writing and drawing are motivated within the fiction. As the analysis has shown, this varies significantly between the two novels: *ATD* explicitly motivates Junior's cartooning as part of his self-understanding and coping strategy, whereas *DWK* motivates only Greg's writing but never his drawings, creating a dissonance that can be read as symptomatic of Greg's general lack of self-awareness and reflective capacity. Technically, real-life diaries do not need to provide any motivation for either writing or drawing, since they are completely private. Yet both novels, through their highly narrative and dialogic orientation, signal an intended public readership, thereby weakening the illusion of privacy and sometimes foregrounding storytelling over self-reflection.

Taken together, these observations demonstrate that cartoon-diaries negotiate reflection and narration in distinctive ways, often privileging narrativity and humor over introspection, and showing reflective moments across verbal and visual modes. This shows how multimodality can reshape the genre conventions of diary novels, both providing ways in which cartoon-diaries deviate from diary novels, but also reinforcing

other conventions. While narrative voice seems mostly unaffected, multimodality does offer a plurality of perspectives, while simultaneously strengthening subjectivity and reader involvement.

7 Conclusion

Based on the genre of cartoon-diaries, this thesis was able to show that multimodality does more than merely illustrate. Multimodality can enrich a text's complexity by, for example, complicating narrative aspects. It destabilizes conventions of internal focalization, distributes subjectivity across modes, and complicates narrative stance, by collapsing the experiencing and narrating diarist temporally and spatially. Multimodality, in these works, is essential for meaning making and goes beyond just illustrative purposes.

Given my professional background as a teacher, as well as the fact that the genre of cartoon-diaries is especially popular in children's and Young Adult literature, as the examples here analyzed exemplify, it seems applicable to address some pedagogical implications.

Scholars have long called for an inclusion of graphic novels and multimodal texts in general in foreign language teaching. Their potential to engage learners as well as support lower-level learners through the use of non-verbal imagery are seen as great potential benefits. Some scholars argue that the inclusion of multimodal texts in EFL classrooms is essential to meeting the demands of 21st century education, as the development of visual and multimodal literacies constitutes a crucial competency (cf. Kachorsky and Serafini). Additionally, having pupils produce multimodal texts in an EFL classroom is vital for developing critical media literacy as well as visual literacy (cf. Alter). Further, multimodal texts can foster intercultural learning and address complex historical or cultural topics (cf. Hankins et al.). In this way, both *ATD* and *DWK* offer considerable potential for use in the EFL classroom. The interaction between verbal and visual modes can lower access barriers to literary texts and enable learners with lower linguistic proficiency to engage meaningfully with literary content. Moreover, as diary novels, both texts lend themselves particularly well to productive

classroom activities, inviting students to create their own multimodal diary entries based on personal experiences. Successful examples thereof have been described by e.g. Jakobsen and Tønnessen. In general, diary novels can facilitate reader identification and encourage empathy and are therefore suitable for classroom readings (Kellner 290).

With regard to *ATD*, this novel lends itself well to classroom discussions of disability and constructions of normality (cf. Crandall; McIntyre and Brown). Often, *ATD* is used to engage students in developing intercultural competences with regard to indigenous identities and experiences (cf. Drege Wågbo; McDougall; Olsen; Szeghi). At the same time, many scholars stress the need for a highly sensitive pedagogical approach, as the novel addresses challenging topics such as alcoholism, violence, and extreme poverty. In this respect, Alexie has been criticized for perpetuating negative stereotypes about Native Americans, however, he consistently argues that these representations reflect real and ongoing issues on contemporary reservations. This means that teaching this novel requires careful contextualization, providing historical context, and close readings with the students to prevent reinforcing or solidifying negative stereotypes or stigmas of Native Americans (see especially Szeghi). As far as the challenging topics are concerned, Ashley argues that the novel provides readers with strategies to confront and cope with real-world challenges, such as poverty, racism, and bullying. According to this perspective, exposing young readers to such content equips them to address these “monsters,” whereas avoidance or censorship does not resolve the underlying issues.

Further, it seems interesting to address the controversies around *ATD*, for it is ranked second on the American Library Association’s list of the ten most frequently challenged books in 2010 due to its perceived vulgarity, including references to masturbation and strong language, as well as anti-Christian content, such as a depiction of Jesus Christ’s flatulence. Concerns over the language and the content have led to its ban in schools across the USA, where it has been removed from curricula or excluded from school libraries (cf. Flagg; Suico). Addressing this aspect in Austrian classrooms might yield a fruitful discussion about the education system in the USA in general.

Finally, it is important to acknowledge a more recent and ethically complex issue superseding pedagogical implications of *ATD* and challenging its use in EFL classrooms. In 2018, there have been allegations of sexual misconduct on part of Sherman Alexie (see for example an article published on March 7th 2018 by Alison Flood in *The Guardian*).⁵⁷ This raises questions about the entanglement of art and its artist, which is particularly relevant when the text is used in educational contexts.⁵⁸ Jeff Spanke, for example, recounts his experience attempting to teach *ATD* in a university YA literature seminar, only to be confronted by student discussions about Alexie's misconduct and its implications for engaging with the novel. Spanke reflects on the difficulty of navigating this tension, ultimately facing the decision if "as teachers, can we find space to honor the integrity of creations while still acknowledging the flaws of the creators?" (Spanke 106–07). This is a question that needs to be answered by teachers individually. If they decide to do a class reading of *ATD*, critical reflection and discussions need to take place, and the students should be involved in this discussion.⁵⁹

Turning to *DWK*, there are also notable pedagogical implications of using this book in a classroom setting. First, there is its extensive transmedia presence, which extends the narrative world far beyond the original books. The series encompasses 20 volumes, as well as several movies with accompanying movie diaries documenting the production process, a musical, and Do-It-Yourself books encouraging children to draw themselves. Despite their visual nature, the books have also been made into successful audiobooks.⁶⁰ The franchise further expands with spin-offs, such as stories told from Rowley's perspective (*Diary Of An Awesome Friendly Kid* by Jeff

⁵⁷ <https://www.theguardian.com/books/2018/mar/07/three-women-go-public-with-sherman-alexie-sexual-harassment-allegations> (Flood)

⁵⁸ As far as the financial ramifications of Alexie's sexual misconduct are concerned, the sales of *ATD* dropped by almost 40% after the allegations and subsequent apology (Milliot).

⁵⁹ A quote that resonated with me personally comes from an online article: "The #MeToo and Time's Up Movements have proven that magnificent things can be created by terrible men; I've already reckoned, many times, with finding out that someone whose work I admire is an abuser in his spare time." (Laban). Unfortunately, the question of these "magnificent things" created by "terrible men" is not an isolated dilemma, but one that must be always negotiated, especially in educational contexts.

⁶⁰ The audiobooks are largely based on the text of the books; however, in certain instances, the illustrations, particularly those containing written elements, are also verbalized (Rickard 3).

Kinney), and an official fan webpage that includes a “For Teachers” section, offering instructional guides and writing prompts. Additionally, the “Fun Zone” provides downloadable worksheets, prompts for drawing, games, and tools to “Wimp Yourself,” creating a personalized *DWK*-style representation of oneself and one’s family, as well as themed emojis (cf. Wimpy Kid Inc.). The series even has its own exhaustive *Diary of a Wimpy Kid Wiki*.⁶¹ Overall, this extensive media ecosystem highlights the series’ pedagogical and creative potential while demonstrating the franchise’s multimodal appeal to engage young audiences in diverse ways.⁶²

However, using *DWK* in EFL classrooms, with a target audience of ca. twelve-year-old children, also requires discussions of masculinity. Greg’s masculinity is fragile, as he is depicted as weak (Potter). The author notes that although the series validates Greg’s inferior masculinity, “the use of humour throughout the text ultimately maintains normative patterns of masculine hegemony” (Potter 4–5). This is evident for example through repeated acts of bullying throughout the novel, where physical violence among males is normalized and patterns of competition and aggression are portrayed as normal features of masculine relationships (Potter 5).⁶³ Bullying is also portrayed as justified to penalize “girly” behavior in boys, which leads to Greg avoiding typically girlish behavior such as sewing or playing with dolls in fear of being bullied for it (Taber and Woloshyn 239). While Taber and Woloshyn see this as problematic, Garbe sees Greg as a good identification source for young boys who suffer from the pressure to be manly by relieving it with humor (Garbe 72–73). Still, because of many bad character traits, Aktaş Kubal and Çetinkaya deem Greg a character “that would be objectionable to be imitated by children” (1111) and they suggest guided classroom discussions so young readers sharpen their perception for the negative messages the book contains.

⁶¹ <https://diary-of-a-wimpy-kid.fandom.com/wiki/Home> (Access: March, 28th 2026)

⁶² Given that the series is likely to continue for the foreseeable future, with the release of the next volume already announced for October 2026, young audiences can remain invested in the characters and storyworld while regularly receiving new content to follow.

⁶³ Greg in *DWK* is both the victim and aggressor of bullying. Older males, like his father and older brother exert power over him, which leads to him bullying weaker males, such as Rowley or his younger brother. Through the use of humor, this is normalized (Potter).

In conclusion, this thesis has shown that multimodality is not merely an aesthetic add-on, but a powerful narrative resource that fundamentally shapes how stories are told and experienced. While cartoon-diaries are primarily associated with children's and YA literature, there seem to be no similar works aimed at an adult audience. Although there are numerous graphic novels with homodiegetic narrators, this specific combination of visual and verbal mode seems to lack from this genre. This gap alone points to promising directions for further research, particularly with regard to how age, audience expectations, and genre conventions interact with multimodal storytelling. Apart from further engaging more systematically with the conventions of children's and YA literature within a multimodal framework, further studies could explore cultural contexts, particularly in the case of *ATD*, in greater depth, which are perspectives that could not be fully developed here due to the scope of this thesis.

Within children's and YA literature, multimodality proves especially effective in heightening reader involvement. This success is reflected not least in the great popularity and commercial success of the works discussed in this thesis. Crucially, multimodality always raises the question of how meaning is distributed across modes and why certain content is conveyed visually rather than verbally and these choices are always intentional. In a media environment that increasingly privileges visual communication, it is hardly surprising that contemporary narratives, especially when they are addressed to younger readers, draw heavily on images to create immediacy, accessibility, and engagement.

As literary forms continue to evolve alongside rapidly changing media landscapes, the study of multimodality remains not only relevant but necessary. New genres will continue to emerge, challenge established categories, and blur boundaries between modes, audiences, and traditions. Examining how these narratives work and why they resonate so strongly will therefore remain an important task for literary research in the years to come.

word count: 32.545

Bibliography

Primary Sources

Alexie, Sherman. *The Absolutely True Diary of a Part-Time Indian*. Andersen Press, 2015

Kinney, Jeff. *Diary of a Wimpy Kid*. Puffin Books, 2015.

Secondary Sources

Abdel-Raheem, Ahmed. "Identity Chains in Newspaper Cartoon Narratives: An Integrative Model." *Journal of Visual Literacy*, vol. 39, no. 1, 2020, pp. 23–48.

Aktaş Kubal, Semra, and Seher Çetinkaya. "The Message Problem in the Children's Book: Diary of a Wimpy Kid." *International Journal of Eurasian Education and Culture*, vol. 6, no. 13, Jan. 2021, pp. 1094–142.

Aldama, Frederick Luis. "Characters in Comic Books." *Characters in Fictional Worlds: Understanding Imaginary Beings in Literature, Film, and Other Media*, De Gruyter, 2010, pp. 318–28. Revision 3.

Al-Jabri, Hanan. "Purification Strategy in the Arabic Translation of Diary of a Wimpy Kid." *International Journal of Arabic-English Studies*, vol. 20, no. 1, 2020, pp. 187–202.

Alós, Julieta. "Paralanguage in the Translation of Children's Graphic Novels into Arabic. Jeff Kinney's Diary of a Wimpy Kid." *Children's Literature in Education*, vol. 2025, no. 56, Nov. 2023, pp. 256–72.

Alter, Grit. "Visual and Media Literacy Put Into Practice: Creating Multimodal Texts in ELT." *Teaching Literature and Language Through Multimodal Texts*, Information Science Reference, 2018, pp. 153–71.

Arend, Helga, and Nina Mahrt. "Gregs Tagebuch, Dork Diaries Und Jede Menge Mangas – Anregungen Zur Leseförderung Anhand von Ausleihzahlen." *Bibliotheksdienst*, vol. 49, no. 9, Sept. 2015, pp. 920–33.

Arfianti, Nur Indah, et al. "Techniques and Accuracy of Ideational Grammatical Metaphor Translation in Diary of Wimpy Kid Novels." *Leksema (Surakarta)*, vol. 3, no. 2, Dec. 2018, pp. 123–35.

Ashley, Stephanie D. *The Absolutely True Diary of a Part-Time Indian: Fighting Monsters*. 2018. San Diego State University.

Baetens, Jan. "Abstraction in Comics." *SubStance*, vol. 40, no. 1, 2011, pp. 94–113.

Bateman, John A., et al. "A Question of Definitions: Foundations for Multimodality: A Response to Charles Forceville's Review." *Visual Communication*, vol. 19, no. 2, 2020, pp. 317–20.

———. *Multimodality: Foundations, Research and Analysis - a Problem-Oriented Introduction*. De Gruyter, 2017.

- . *Text and Image: A Critical Introduction to the Visual/Verbal Divide*. Routledge, 2014.
- Bauer, Iris. *The Diary Novel in English*. 2003. Universität Wien.
- Bodden, Valerie. “Navigating Race and Identity.” *Young Adult*, ABDO Publishing Company, 2016.
- Bouzonviller, Elisabeth. “Laughing Together as a Strategy of Survivance for Native American Communities.” *The Concept of Community from a Global Perspective*, Brill, 2024, pp. 207–25.
- Chatman, Seymour. *Story and Discourse : Narrative Structure in Fiction and Film*. Cornell University Press, 1980.
- Cohn, Dorrit. *Transparent Minds. Narrative Modes for Presenting Consciousness in Fiction*. Princeton University Press, 1983.
- Crandall, Bryan R. “Adding a Disability Perspective When Reading Adolescent Literature: Sherman Alexie’s *The Absolutely True Diary of a Part-Time Indian*.” *The ALAN Review*, vol. 36, no. 2, 2009, pp. 71–78.
- Doğru, Büşra. *The Concept of “Homo Sacer” in Sherman Alexie’s Reservation Blues, The Lone Ranger and Tonto Fistfight in Heaven and The Absolutely True Diary of a Part-Time Indian*. 2022. Dokuz Eylül University.
- Doloughan, Fiona J. *Contemporary Narrative : Textual Production, Multimodality and Multiliteracies*. Continuum, 2011.
- Drege Wågbø, Ole Martin. *Encouraging Empathy through The Absolutely True Diary of a Part-Time Indian*. 2022. Norwegian University of Science and Technology, Bachelor thesis.
- Eisner, Will. *Comics and Sequential Art: Principles and Practices from the Legendary Cartoonist*. 1985. Rev. ed., W. W. Norton & Company, Incorporated, 2008.
- Elstad, Martine. *Multimodal Young Adult Fiction and Reading Motivation. An Analysis of Sherman Alexie’s The Absolutely True Diary of a Part-Time Indian and Mark Haddon’s The Curious Incident of the Dog in the Night-Time and Their Possible Influence on Reading Motivation among Young Male Readers*. 2021. The Arctic University of Norway.
- Evans, Janet. “From Comics, Graphic Novels and Picturebooks to Fusion Texts: A New Kid on the Block!” *Education*, vol. 3, no. 13, 2013, pp. 233–48.
- Field, Trevor. *Form and Function in the Diary Novel*. Barnes & Noble, 1989.
- Flagg, Gordon. “Concerned Parents School Educators About Reading Choices.” *American Libraries*, vol. 43, no. 1, Jan. 2012, p. 16.
- Flood, Alison. “Three Women Go Public with Sherman Alexie Sexual Harassment Allegations.” *The Guardian*, 7 Mar. 2018. *The Guardian*, <https://www.theguardian.com/books/2018/mar/07/three-women-go-public-with-sherman-alexie-sexual-harassment-allegations>.

- Forney, Ellen. "Interview." *The Absolutely True Diary of a Part-Time Indian*, Andersen Press, 2015, pp. 239–45.
- Garbe, Christine. "Antiheldengeschichten." *Attraktive Lesestoffe (Nicht Nur) Für Jungen. Erzählmuster Und Beispielanalysen Zu Populärer Kinder- Und Jugendliteratur*, Schneider Verlag Hohengehren, 2018, pp. 68–84.
- Garić, Vanja Vukićević. "Part-Time Identities and Full-Time Narration as an Absolution in Alexie's *The Absolutely True Diary of a Part-Time Indian*." *Neohelicon*, vol. 44, 2017, pp. 189–98.
- Gaudreault, André. "Narration and Monstration in the Cinema." *Journal of Film and Video*, vol. 39, no. 2, 1987, pp. 29–36.
- Genette, Gérard. *Narrative Discourse*. Paperback ed., Repr, Basil Blackwell, 1986.
- Gleason, Carrie. "Bridging the Divide Between Non-Readers and Lifelong Readers with Hi-Lo Books." *American Association of School Librarians Knowledge Quest*, vol. 3, no. 4, Mar. 2011, pp. 22–26.
- Goldstein, David S. "Sacred Hoop Dreams: Basketball in the Work of Sherman Alexie." *Ethnic Studies Review*, vol. 32, no. 1, 2009, pp. 77–88.
- Groensteen, Thierry. *Comics and Narration*. Translated by Ann Miller, University Press of Mississippi, 2013.
- . *System of Comics*. Translated by Bart Beaty and Nick Nguyen, University Press of Mississippi, 2007.
- Gruber, Eva. *Humor in Contemporary Native North American Literature . Reimagining Nativeness*. Boydell & Brewer, 2008.
- Haller, Christina. *Children's Literature - a Translation Analysis of Jeff Kinney's Diary of a Wimpy Kid*. 2014. Karl-Franzens-Universität Graz.
- Hankins, Bonnie, et al. "Exploring the Graphic Novel's Popularity and Advocating for Its Integration in the World-Language Classroom." *Bookbird*, vol. 62, no. 4, Jan. 2024, pp. 62–66.
- Heritage, Stuart. "Wimpy Kid Author Jeff Kinney: 'I've Sold 300m Books. What's next?'" *The Guardian*, 1 Feb. 2026. Books. *The Guardian*, <https://www.theguardian.com/books/2026/feb/01/wimpy-kid-author-jeff-kinney-ive-sold-300m-books-whats-next>.
- Horstkotte, Silke, and Nancy Pedri. "Focalization in Graphic Narrative." *Narrative*, vol. 19, no. 3, 2011, pp. 330–57.
- Hughes, Janette Michelle, et al. "Adolescents and 'Autographics': Reading and Writing Coming-of-Age Graphic Novels." *Adolescent & Adult Literacy*, vol. 54, no. 8, 2011, pp. 601–12.
- Hui, Haifeng. "Currency's Reversed Marginal Role in Children's Literature: Loans, Debts, Mum Bucks and Their Subversion in 'Diary of a Wimpy Kid.'" *Literature Compass*, online, vol. 18, no. 3, 2021.

- . “Narrative Voice, Mind-Reading and Empathy in Children’s Diaries: The Dominant Autodiegetic Voice in *Diary of a Wimpy Kid*.” *Journal of Narrative Theory*, vol. 4, no. 3, Sept. 2024, pp. 395–415.
- Ibarrola Armendariz, Aitor. “Sherman Alexie’s *The Toughest Indian in the World*. Amalgamating the Oral and the Written Traditions.” *Complutense Journal of English Studies (Internet)*, vol. 23, 2015, pp. 137–60.
- Jakobsen, Ingrid Karoline, and Elise Seip Tønnessen. “A Design-Oriented Analysis of Multimodality in English as a Foreign Language.” *Designs for Learning*, vol. 10, no. 1, 2018, pp. 40–52.
- Jannetta, Armando E. *Ethnopoetics of the Minority Voice. An Introduction to the Politics of Dialogism and Difference in Métis Literature*. Wißner, 2001.
- Johnson, Jan. “Healing the Soul Wound in Flight and The Absolutely True Diary of a Part-Time Indian.” *Sherman Alexie: A Collection of Critical Essays*, University of Utah Press, 2010, pp. 224–40.
- Kachorsky, Dani, and Frank Serafini. “From Picturebooks to Propaganda: Developing Visual and Multimodal Literacies.” *Teaching Literature and Language Through Multimodal Texts*, Information Science Reference, 2018, pp. 70–92.
- Katanski, Amelia V. “School Stories: Education, Coming of Age, and the American Indian Literary Tradition.” *Coming of Age*, Salem Press, 2013, pp. 265–81.
- Kellner, Renate. *Der Tagebuchroman Als Literarische Gattung : Thematologische, Poetologische Und Narratologische Aspekte*. De Gruyter, 2015.
- Kertzer, Adrienne. “Not Exactly: Intertextual Identities and Risky Laughter in Sherman Alexie’s *The Absolutely True Diary of a Part-Time Indian*.” *Children’s Literature*, vol. 40, no. 1, 2012, pp. 49–77.
- Kress, Gunther R. *Literacy in the New Media Age*. Routledge, 2003.
- Kress, Gunther R., and Theo van Leeuwen. *Multimodal Discourse : The Modes and Media of Contemporary Communication*. Hodder Education, 2001.
- . *Reading Images : The Grammar of Visual Design*. Routledge, 2006.
- Kukkonen, Karin. “Comics as a Test Case for Transmedial Narratology.” *SubStance*, vol. 40, no. 1, 2011, pp. 34–52. 124.
- . *Contemporary Comics Storytelling*. University of Nebraska Press, 2013.
- Kurnia, Nandy Intan, et al. “The Battle of Domination between Adults and Children in Jeff Kinney’s *Diary of a Wimpy Kid*.” *Jordan Journal of Modern Languages and Literature*, vol. 16, no. 1, pp. 199–215.
- Laban, Monique. “Why Sherman Alexie’s Sexual Misconduct Feels Like a Betrayal.” *Electric Literature*, 03 2018, <https://electricliterature.com/why-sherman-alexies-sexual-misconduct-feels-like-a-betrayal/>.
- Lacy, Meagan. “Portraits of Children of Alcoholics: Stories That Add Hope to Hope.” *Children’s Literature in Education*, vol. 46, no. 4, 2015, pp. 343–58.

- Lemke, Jay. "Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text." *Reading Science: Critical and Functional Perspectives on Discourses of Science*, Routledge, 1998, pp. 87–113.
- Lewis, David K. *Convention. A Philosophical Study*. Harvard University Press, 1969.
- MacIntyre, Pamela. "The Rise of the Illustrated YA Novel: Challenges to Form and Ideology." *The International Journal of the Book*, vol. 8, no. 1, 2011, pp. 135–44.
- Martens, Lorna. *The Diary Novel*. Cambridge University Press, 1985.
- McCloud, Scott. *Understanding Comics. The Invisible Art*. Reprint, William Morrow, an imprint of Harper Collins Publishers, 1993.
- McDougall, Morgan E. *Teaching Native American and Middle East American Literature in the Secondary School Classroom*. 2018. Graduate College of Bowling Green State University.
- McIntyre, Siobhan, and Alan Brown. "Dare to Be Different: Celebrating Difference and Redefining Disability in Sherman Alexie's *The Absolutely True Diary of a Part-Time Indian*." *The ALAN Review*, vol. 44, no. 1, Oct. 2016, pp. 79–85.
- McLuhan, Marshall. *Understanding Media: The Extensions of Man*. 1964. Cambridge, Mass., 1994.
- Mikkonen, Kai. *The Narratology of Comic Art*. Routledge, 2017.
- Milliot, Jim. "Taking the Measure of Sexual Misconduct Charges in Publishing." *Publishers Weekly*, vol. 265, no. 17, 2018, p. online.
- Mitchell, W. J. T. "Showing Seeing: A Critique of Visual Culture." *Journal of Visual Culture*, vol. 1, no. 2, 2002, pp. 165–81.
- . "There Are No Visual Media." *Journal of Visual Culture*, vol. 4, no. 2, 2005, pp. 257–66.
- Moon, Set-Byul. "Superheroes Do Not Live on the Rez: The Nomadic Identity for Native Indian Young Adults in Sherman Alexie's *The Absolutely True Diary of a Part-Time Indian*." *Trans-Humanities Journal*, vol. 9, no. 1, 2016, pp. 151–70.
- Moss, Marissa. *Amelia's Notebook*. <https://www.marissamoss.com/books/amelias-notebook/>. Accessed 14 Apr. 2026.
- "Most Printed Fictional Diary Series." *Guinness World Records*, <https://www.guinnessworldrecords.com/world-records/397196-most-printed-fictional-diary-series.html>. Accessed 28 Mar. 2026.
- Mroczek-Bayci, Emily. "Show Me The Funny: Humor Keeps Tweens And Teens Reading. Why Isn't There More?" *School Library Journal*, vol. 86, no. 6, June 2022, pp. 30–34.
- Nazari Robati, Fatemeh Zahra, and Fatemeh Zand. "Translation of Taboos: The Absolutely True Diary of a Part-Time Indian." *International Journal of Applied Linguistics & English Literature*, vol. 7, no. 3, 2018, pp. 35–40.

- Olsen, Andrea. *How Can The Absolutely True Diary of a Part-Time Indian and Reading Logs Be Used as Teaching Tools to Develop and Demonstrate Pupils' Intercultural Competence?* 2022. Høgskolen i Østfold.
- Oxford Learner's Dictionary. "Image." *Oxford Learner's Dictionary*, <https://www.oxford-learnersdictionaries.com/definition/english/image?q=image>. Accessed 18 Apr. 2026.
- Packard, Stephan, et al. *Comicanalyse. Eine Einführung*. J.B. Metzler, 2019.
- Page, Ruth. *New Perspectives on Narrative and Multimodality*. Routledge, 2010.
- Palmer, Joye Tompkins. *Identity Formation in Sherman Alexie's The Lone Ranger and Tonto Fistfight in Heaven, Flight and The Absolutely True Diary of a Part-Time Indian*. 2017. University of North Carolina.
- Parsegova, Gayana. *Defining the Bildungsroman Genre in Youth Literature: Utilizing Young Adult Texts to Bridge the Gap of Cultural Duality with The House on Mango Street, The Absolutely True Diary of a Part-Time Indian, and Breaking Stalin's Nose*. 2020. San Diego State University.
- Pischedda, Pier Simone. "A Corpus-Based Study on the Translation of English Ideophones in Italian Picture Books: The Case of the Diary of a Wimpy Kid." *Languages (Basel)*, vol. 7, no. 3, 2022, pp. 224–38.
- Potter, Troy. "Comedy and Masculine Hegemony: Adapting Masculinities in Humorous Children's Fiction." *Children's Literature in Education*, online, 2024.
- Prince, Gerald. "The Diary Novel: Notes for the Definition of a Sub-Genre." *Neophilologus*, no. 59, 1975, pp. 477–81. 4.
- Prodanović Stankić, Diana, and Helga Begonja. "Expressive Language in Translation of Books for Children: A Corpus-Based Study." *Círculo de Lingüística Aplicada a La Comunicación*, vol. 100, no. 100, 2024, pp. 235–43.
- Reid, Stephanie F., and Frank Serafini. "More Than Words: An Investigation of the Middle-Grade Multimodal Novel." *Journal of Children's Literature*, vol. 44, no. 2, 2018, pp. 32–44.
- Rickard, Rachel L. *Wimps, Dorks, and Reluctant Readers: Redefining Literacy in Multimodal Middle Grade Diary Books*. 2014. Eastern Michigan University.
- Ryan, Marie-Laure. "Toward a Definition of Narrative." *The Cambridge Companion to Narrative*, Cambridge University Press, 2007.
- Santoso, Heri Dwi, and Dwi Ario Fajar. "Greg Heffley's Role Confusion in the Diary of a Wimpy Kid (2007): A Psychosocial Analysis." *Lire Journal*, vol. 7, no. 1, May 2023, pp. 140–50.
- Schlaffer, Tanja. *North American Na(Rra)Tives. De/Constructing the Other in Selected Native American Young Adult Literature*. 2018. Universität Wien.
- Sousanis, Nick. *Unflattening*. Harvard University Press, 2015.

- Spanke, Jeff. "Magnificent Things and Terrible Men: Teaching Sherman Alexie in the Age of #MeToo." *English Education*, vol. 51, no. 1, 2018, pp. 101–08.
- Sparks, Andrea. "'Lost, Squared' Reservation Realism and the Borderless Imaginary." *Critical Explorations of Young Adult Literature*, Routledge, 2020, pp. 67–81.
- Spring, Erin. "'I Think I Was Born with a Suitcase' Blackfoot Adolescent Readers' Responses to Sherman Alexie's *The Absolutely True Diary of a Part-Time Indian*." *Young People Reading. Empirical Research Across International Contexts*, Routledge, 2018, pp. 106–20.
- Stöckl, Hartmut. "Linguistic Multimodality - Multimodal Linguistics: A State-of-the-Art Sketch." *Multimodality: Disciplinary Thoughts and the Challenge of Diversity*, De Gruyter, 2019.
- Suhr-Sytsma, Mandy. "Indigenous School Stories: Alternatives to Multiculturalism in Sherman Alexie's *The Absolutely True Diary of a Part-Time Indian* and Joseph Bruchac's *The Heart of a Chief*." *Self-Determined Stories: The Indigenous Reinvention of Young Adult Literature*, Michigan State University Press, 2019, pp. 25–64.
- Suico, Terri. "A Review of Defending Frequently Challenged Young Adult Books, Teaching Banned Books, and Books Under Fire." *Study and Scrutiny*, vol. 6, no. 1, Aug. 2023, pp. 197–204.
- Szeghi, Tereza M. "The Possibilities and Pitfalls in Teaching Sherman Alexie's *The Absolutely True Diary of a Part-Time Indian*." *Multiethnic American Literatures.*, McFarland & Company, Inc., Publishers, 2015, pp. 72–99.
- Taber, Nancy, and Vera Woloshyn. "Dumb Dorky Girls and Wimpy Boys: Gendered Themes in Diary Cartoon Novels." *Children's Literature in Education*, vol. 42, 2011, pp. 226–42.
- Teuton, Christopher B. *Deep Waters : The Textual Continuum in American Indian Literature*. University of Nebraska Press, 2010.
- Thon, Jan-Noel. "Comics Narratology." *Handbook of Comics and Graphic Narratives*, vol. 11, De Gruyter, 2021, pp. 99–120.
- Unterweger, Julia. *'I Just Want to Not Be Me' : The Struggles of Growing up in Young Adult Literature*. 2017. Karl-Franzens-Universität Graz.
- Verusha Ali, Helena, et al. "A Study of Presupposition and Conversational Implicatures in the Comic Strip of 'The Diary of a Wimpy Kid.'" *Jurnal Onoma: Pendidikan, Bahasa, Dan Sastra*, vol. 9, no. 2, Oct. 2023, pp. 1353–64.
- Vogel, Mark. "Half Child/Half Adult: Sherman Alexie's Hybrid Young Adult Fiction." *Sherman Alexie. Critical Insights*, Salem Press, 2012, pp. 106–25.
- Voinea, Mihaela, and Mariana Norel. "For a Pedagogy of Reading in the Digital/ Post-modern Age. *Diary of a Wimpy Kid* – Invitation to a 'Different' Reading." *Philologica Jassyensia*, vol. 24, no. 2, July 2016, pp. 317–25.

- Wahpeconiah, Tammy. "Navigating the River of the World: Collective Trauma in The Absolutely True Diary of a Part-Time Indian." *Sherman Alexie. Critical Insights*, Salem Press, 2012, pp. 35–52.
- Walton, Kendall L. *Mimesis as Make-Believe : On the Foundations of the Representational Arts*. Harvard University Press, 1990.
- Wayland, Nerida. "Representations of Happiness in Comedic Young Adult Fiction: Happy Are the Wretched." *Jeunesse: Young People, Texts, Cultures*, vol. 7, no. 2, 2015, pp. 86–106.
- Wildfeuer, Janina, et al. *Multimodality: Disciplinary Thoughts and the Challenge of Diversity*. De Gruyter, 2019.
- Wimpy Kid Inc. "Wimpy Kid · Official Website for Jeff Kinney's Diary of a Wimpy Kid." *Wimpy Kid*, <https://wimpykid.com/>. Accessed 28 Mar. 2026.
- Wolf, Werner. "Transmedial Narratology: Theoretical Foundations and Some Applications (Fiction, Single Pictures, Instrumental Music)." *Narrative*, vol. 25, no. 3, Oct. 2017, pp. 256–85.
- Wyle, Andrea. "Expanding the View of First-Person Narration." *Children's Literature in Education*, vol. 30, no. 3, 1999, pp. 185–202.

Cartoon-Diaries – Further Reading

These works are mentioned in this thesis.

- Kinney, Jeff. *Diary of an Awesome Friendly Kid. Rowley Jefferson's Journal*. Hachette Book Group, 2019.
- Kinney, Jeff. *Diary of a Wimpy Kid. Partypoooper*. Puffin Books, 2025.
- Moss, Marissa. *Amelia's Notebook*. Tricycle Press, 1995.
- Russel, Rachel Renée. *Dork Diaries. Tales from a Not-So Fabulous Life*. Simon & Schuster, 2009.

Appendix

AI Acknowledgment

“A fool with a tool is still a fool”

I include this quote to emphasize that tools do not replace independent thought or academic responsibility. While working on this thesis, I remained fully aware that any digital tool can only support, but never substitute, critical reflection, careful reading, and original analysis.

This project started in February 2021, prior to the widespread integration of generative AI tools into academic practice. The majority of the research and writing was completed independently before such tools became commonly accessible. At a later stage of my research phase, I used the tool *NotebookLM* (<https://notebooklm.google/>) to gain an overview of a large body of secondary literature related to *DWK* and *ATD* in general. I used the tool primarily to generate summaries of texts in order to identify potentially relevant arguments and passages. I subsequently read and evaluated the relevant texts, and any integration of ideas into this thesis is based on my own reading and analysis. In addition, I used the free version of *ChatGPT* (<https://chatgpt.com/>) to receive feedback on the formulation of the thesis title, research question, and (sub-)headings, and incorporated these refinements. I also used *ChatGPT* to rephrase sections I had written in order to improve clarity and academic style. These prompts were based exclusively on my own ideas, research, and written drafts. I reviewed the output and used it as a suggestion, instead of copying it verbatim, in order to maintain a consistent personal academic voice. At no point did I use AI tools to generate original arguments, interpretations, or research findings. The intellectual work, analysis, and conclusions presented in this thesis are entirely my own. AI tools were used solely for purposes of language refinement, structural feedback, and literature overview. I take full responsibility for the content of this thesis.

Records of the relevant interactions with AI tools have been retained and can be provided if required.

Abstract

This thesis analyzes the narratological implications of multimodality in the genre of cartoon-diaries based on the two works *Diary of a Wimpy Kid* (2007) and *The Absolutely True Diary of a Part-Time Indian* (2007). These novels are highly popular examples of this genre and have been widely discussed in research, although never from a genre-based multimodal perspective. In addition to multimodality as a framework, the analysis draws on narratological categories developed by the French literary theorist Gérard Genette, in particular in relation to narrative stance, understood through the interplay of narrative voice, focalization, and subjectivity, as well as their specific configuration in monomodal diary narration. With regard to their multimodality, cartoon-diaries are constituted by the interaction of verbal written language and cartoony imagery. On this basis, a close reading of the two novels was conducted within a genre-sensitive comparative framework. The analysis shows that multimodality significantly shapes the genre, particularly by opening up an interplay between internal and external perspectives, when the diarist maintains an autodiegetic voice in the running text and visually constructs him:herself in the cartoons. Thus, narrative stance and its characteristic immediacy typical of monomodal diary narration appear more dynamic in the multimodal format. Overall, these effects can be traced back to the multimodal interaction between telling and showing, as well as the spatial and temporal organization of the multimodal text. This thesis demonstrates that multimodality is not merely a form of illustration in cartoon-diaries, but a constitutive element that reshapes their narratological structure and meaning-making processes.

Abstract [Deutsch]

Diese Masterarbeit analysiert den Einfluss von Multimodalität auf die Narratologie von Cartoon-Tagebüchern anhand der Romane *Diary of a Wimpy Kid* (2007) [*Gregs Tagebuch – Von Idioten umzingelt*] sowie *The Absolutely True Diary of a Part-Time Indian* (2007) [*Das absolut wahre Tagebuch eines Teilzeit Indianers*]. Beide Werke erfreuen sich großer Popularität und wurden bereits umfangreich wissenschaftlich untersucht, jedoch bislang nicht aus einer genrebezogenen multimodalen Perspektive. Neben der Multimodalität als methodologischem Rahmen stützt sich die Analyse auf das narratologische Konzept des französischen Literaturwissenschaftlers Gérard Genette, insbesondere im Hinblick auf die Erzählhaltung, verstanden als Zusammenspiel von Erzählstimme, Fokalisierung und Subjektivität, sowie deren konkrete Ausgestaltung in monomodalen Tagebuchromanen. Die Multimodalität in beiden Texten ergibt sich aus der Interaktion von geschriebener Sprache und cartoonspezifischen Zeichnungen. Auf dieser Grundlage erfolgt ein genrebezogenes vergleichendes Close Reading der beiden Romane. Die Analyse zeigt, dass Multimodalität das Genre signifikant prägt, insbesondere durch das Zusammenspiel innerer und äußerer Perspektiven, wenn die Tagebuchschreibenden im Fließtext eine autodiegetische Perspektive einnehmen und sich selbst zeichnerisch darstellen. Die konventionelle Erzählhaltung von Tagebuchromanen und ihre typische Unmittelbarkeit erweisen sich letztendlich als besonders dynamisch im multimodalen Kontext. Insgesamt lassen sich diese Effekte auf die multimodale Interaktion von berichtendem Erzählen und szenischem Darstellen sowie auf die spatio-temporalen Strukturierungen der Texte zurückführen. Die Arbeit zeigt somit, dass Multimodalität über eine rein illustrative Funktion hinausgeht und als konstitutives Element die narratologischen Strukturen und Bedeutungskonstruktion dieses Genres wesentlich mitgestaltet.