

# Cognitive Linguistics Research

---

**Editors**

Dirk Geeraerts  
Dagmar Divjak  
John R. Taylor

**Honorary editors**

René Dirven  
Ronald W. Langacker

**Volume 55**

# **Viewpoint and the Fabric of Meaning**

---

Form and Use of Viewpoint Tools  
across Languages and Modalities

Edited by  
Barbara Dancygier  
Wei-lun Lu  
Arie Verhagen



**DE GRUYTER**  
MOUTON

ISBN 978-3-11-036907-6  
e-ISBN (PDF) 978-3-11-036546-7  
e-ISBN (EPUB) 978-3-11-039307-1  
ISSN 1861-4132

**Library of Congress Cataloging-in-Publication Data**

A CIP catalog record for this book has been applied for at the Library of Congress.

**Bibliographic information published by the Deutsche Nationalbibliothek**

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;  
detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2016 Walter de Gruyter GmbH, Berlin/Boston  
Typesetting: PTP-Berlin, Protago-T<sub>E</sub>X-Production GmbH, Berlin  
Printing and binding: CPI books GmbH, Leck  
♾ Printed on acid-free paper  
Printed in Germany

[www.degruyter.com](http://www.degruyter.com)



# Table of contents

List of contributors — vii

Arie Verhagen

**Introduction: On tools for weaving meaning out of viewpoint threads** — 1

## Part I: The ubiquity of viewpoint

Barbara Dancygier and Lieven Vandelanotte

**Discourse viewpoint as network** — 13

Jeroen Vanderbiesen

**Mixed viewpoints and the quotative-reportive cline in German: Reported speech and reportive evidentiality** — 41

Katsunobu Izutsu and Mitsuko Narita Izutsu

**Viewpoint fusion for realism enhancement in Ainu and Japanese narratives** — 93

Elisabeth Engberg-Pedersen and Ditte Boeg Thomsen

**The socio-cognitive foundation of Danish perspective-mixing dialogue particles** — 125

## Part II: Across languages

Kobie van Krieken, José Sanders, Hans Hoeken

**Blended viewpoints, mediated witnesses: A cognitive linguistic approach to news narratives** — 145

Wei-lun Lu and Arie Verhagen

**Shifting viewpoints: How does that actually work across languages? An exercise in parallel text analysis** — 169

Ad Foolen and Toshiko Yamaguchi

**Perspective: Kawabata's *Beauty and Sadness* and its translations into English, German, and Dutch** — 191

## Part III: Across modalities

Chie Fukada

**The dynamic interplay between words and pictures in picture storybooks: How visual and verbal information interact and affect the readers' viewpoint and understanding** — 217

Eve Sweetser and Kashmiri Stec

**Maintaining multiple viewpoints with gaze** — 237

Maria Josep Jarque and Esther Pascual

**Mixed viewpoints in factual and fictive discourse in Catalan Sign Language narratives** — 259

Barbara Dancygier

**Concluding remarks: Why viewpoint matters** — 281

**Index** — 289

Chie Fukada

# The dynamic interplay between words and pictures in picture storybooks: How visual and verbal information interact and affect the readers' viewpoint and understanding

**Abstract:** Like language, visual images have a “grammar”, a set of elements and rules for producing and understanding meanings when they are used in social communication (Kress and van Leewen 2006). Stories can therefore appear not only in narrative forms but also in visual forms, and their combination can afford a better understanding of the story represented. Picture storybooks present fictive worlds through this combination, but the viewpoint reflected in verbal narratives, which is relatively stable and objective (Matsuoka 1987), is not always consistent with that of the visuals, producing “mixed” points of view. However, readers can create consistent stories by employing their everyday experiences and prior knowledge to adjust these different viewpoints. The current study conducts an in-depth analysis of a picture storybook at both verbal and visual levels of representation, and discusses how the three types of viewpoints, i.e. the viewpoints in the narrative and the visuals and that of the readers, are integrated to produce a coherent story in the readers’ minds. It elucidates the mechanisms of how picture storybooks prompt readers to set up, understand, and become involved in the fictive worlds expressed in them.

---

**Note:** This study is an expansion and explication of Fukada (2011). It was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), 2011–2014 (24520541, Kazumi Taniguchi). I would like to thank the anonymous reviewer for their comments and suggestions for improving my paper. I would also like to express my gratitude to Takatsugu Kojima, Haruhiko Yamaguchi, José Sanders, Eve Sweetser, Wei-lun Lu, Sandra Healy and Yasushi Tsubota for their insightful comments and suggestions and to Kathleen Yamane for her careful reading of my first draft. I am also indebted to Ryoma Nishizawa, Fumiko Kishimoto, Yuko Suzuki and Ami Tsukamoto for their assistance. All remaining errors are, of course, my own.

## 1 Introduction

Picture storybooks present fictive worlds through the interplay between two different modes of representation: visual and verbal. The readers integrate these two kinds of information into a coherent story by associating them with their own knowledge and everyday experience. Here, two questions arise: (i) In what ways do these two different sources of information merge to create a coherent story? and (ii) What kinds of words and/or pictures evoke the readers' everyday experiences and enable them to understand the story?

Although there is extensive literature on picture books in the fields of children's literature and literacy education, much of the work deals with questions of literacy such as "What ways of reading are effective for children to understand the story more deeply?" (e.g., Nakamura 1995) or "How are the verbal, visual, oral performance, and instructional cues intertwined to promote children's interpretation of the story?" (e.g., Golden and Gerber 1990). Nikolajeva and Scott (2001) provide one of the few exceptions in that their study addresses the dynamics between words and images in the genre of picture books, but their claims need to be validated or supported by experimental evidence. Another exception is the work of Johnston (2012), which contains an analysis of a picture storybook with no words. However, she mainly focuses on how the story grows out of the images.

Furthermore, a number of studies within the framework of mental spaces have provided a detailed account of our mental processes of understanding literary texts or visual information like gestures and sign language (e.g., Fauconnier and Sweetser 1996 and Dancygier and Sweetser 2012), but no research has been conducted on picture books in this field so far. As for the interplay between visual and verbal information, some studies in cognitive science and artificial intelligence have addressed the issue of how both types of information are processed and integrated in learning language (e.g., Oka et al. 2013), but very few have dealt with the dynamics of text, images and the reader's understanding process of picture books.

In this study, therefore, I conduct an in-depth and practical analysis of the picture storybook *Shiroi Usagi to Kuroi Usagi* (hereinafter, simply *Usagi*) at each level of representation, i.e. the visual and the verbal, and explore what kind of pictures and/or words affect the reader's viewpoint, understanding, or even involvement, and how they do this. Section 2 provides a brief commentary on *Usagi*, and Section 3 investigates the illustrations and text of the book in detail, revealing how they differ and how they combine. Section 4 presents an extended discussion of the results of our experimental pilot study conducted in 2013. Finally, Section 5 gives some concluding remarks and addresses remaining issues.

## 2 The Object of investigation: *Usagi*

The picture book treated in this study, *Usagi*, is the Japanese version of *The Rabbits' Wedding* (hereinafter, *TRW*), a popular illustrated children's book by Garth Williams published in America in 1958. The book has been read in Japan for over 50 years, and thus it has become familiar to Japanese people of many ages.<sup>1</sup> Matsuoka (1987: 51–56) proposes that the picture storybooks which have been read for over 25 years should be classified as outstanding (although she does not mention what characteristics those books share), and so by this standard the book *Usagi* certainly deserves deeper investigation.

Although the title is not a direct translation of *TRW* and some names of the games and plants were changed to those that Japanese readers can easily imagine, the story develops in the same way as in the original English version.<sup>2</sup> A synopsis of *TRW* is as follows:

Two little rabbits, one white and the other black, played together happily in the forest. But in between the games of Hop Skip And Jump Me and Race Around The Blackberry Bush the black rabbit stopped and sighed. "I'm just thinking," he would say, when the white rabbit asked him what was the matter. But he finally admitted he was wishing – wishing that he and the white rabbit could be together forever and always. And after he had wished a little harder his wish came true.

(quoted from the jacket of *TRW*)

As with the vast majority of picture books, the relationship between words and pictures is, at least at first glance, symmetrical, consonant, or complementary, although upon closer examination even books of this kind contain some interesting contrasts or discrepancies between them (see discussion in Nikolajeva and Scott 2001: 14). The next sections, in conducting an in-depth analysis of the picture book *Usagi* at both the visual and verbal levels of information, not only show such contrasts but also discuss how they are adjusted by the reader to create a coherent single story.

---

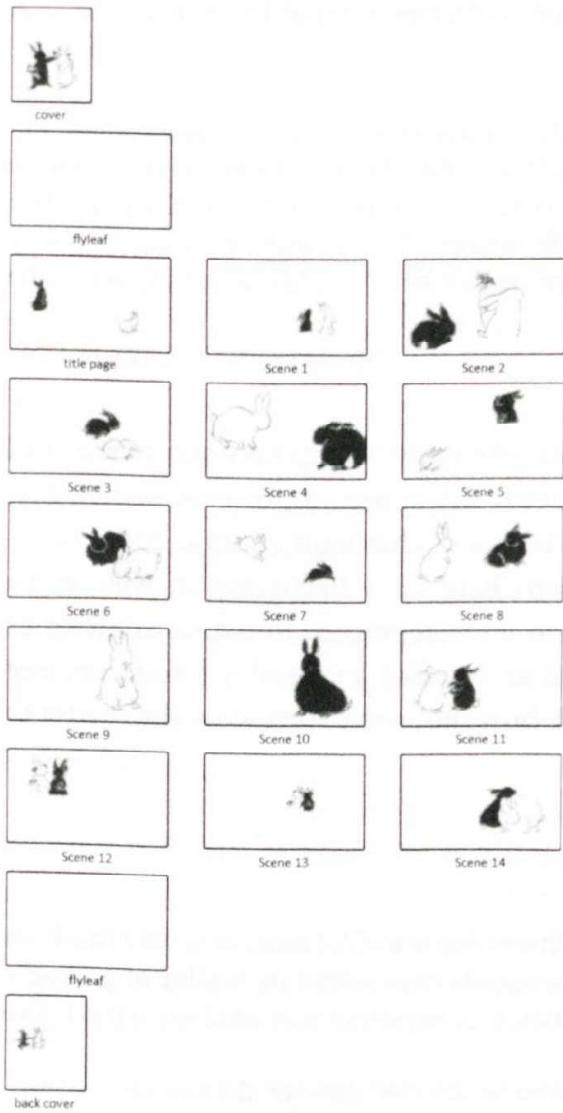
<sup>1</sup> I am also very familiar with the book *Usagi*. When I was a child, I repeatedly read this book with my father and by myself; furthermore, I personally experienced the feeling of gradually coming to understand the feelings of the two rabbits, thus becoming more involved in this fictive world.

<sup>2</sup> The following differences in bookmaking can also be observed between the English and Japanese versions: (i) only the English version has a double-spread title page with no picture after the flyleaf; and (ii) the closing remarks are written on the back flyleaf with no pictures in the English version, while they are incorporated into the final page of the story in the Japanese version.

### 3 An analysis of *Usagi*: Contrasts and harmonies between pictures and words

#### 3.1 What the pictures represent

In this section, I analyse the pictures in *Usagi* with special attention to the colours used in this book, the size of the two characters, their facial expressions, and their eye (or face or body) orientations. Although Kress and van Leeuwen (2006) have concentrated on explicating the grammar in visual communication in



**Figure 1:** Line drawings of the picture storybook *Usagi*

western cultures, not in non-western cultures including Japan, and they have not dealt with the sequence of pictures in picture books, the analysis in this paper is based largely on their findings of how the depicted elements and their structures (colour, size, framing, etc.) of visual design are combined into meaningful wholes. Before analysing the pictures in *Usagi*, I will show how the story is depicted in its entirety in Figure 1. Figure 1 presents rough line drawings of the two rabbits in all of the pictures in the book *Usagi*, created for the discussion below.

### 3.1.1 The use of colour

The colours used in *Usagi* are white (mostly for the white rabbit), black (mostly for the black rabbit), and yellow (mostly for the background including plants). The yellow in most scenes is subdued and greenish, although a pure yellow is used for the flowers, and the background of Scene 9. Combined with the soft feel of the brush, this subdued or low-saturation colour throughout the book gives the story a tranquil, calm atmosphere. The colour can also be viewed as representing the subtle or tender characteristics of the black rabbit (see discussions in Kress and van Leeuwen 2006: 234), while the pure yellow in Scene 9 can be seen as reflecting the feeling of the white rabbit, namely a complete but delightful surprise at the words of the black rabbit.

### 3.1.2 Changes in the size of the characters

As shown in Figure 1, the sizes of the characters depicted change from scene to scene. The repetition of large and small sizes can be seen in the scenes leading up to the climax (i.e., Scenes 9 and 10) of the story.<sup>3</sup> At the climax, the two rabbits are depicted largest, and after that (more precisely, from Scenes 11 to 13) they get smaller. As Kress and van Leeuwen (2006: 124–129) argue, the choice between close-up, medium shot, long shot, etc. indicates, by analogy with everyday social interaction, the distance of the represented participants from the viewers: at intimate distance we only see the face or head of the other person, at close personal distance we see his or her head and shoulders, at far personal distance we see him or her from the waist up, at close social distance we see his or her whole figure, at far social distance we can see the whole figure with space around it, and at public

---

<sup>3</sup> Strictly speaking, at the text level, the climax begins with the second dialogue between the protagonists in Scene 8 and ends in Scene 11 with the white rabbit's acceptance of the black rabbit's marriage proposal.

distance we can see his or her whole body with a lot of space around him or her. Although this distinction is not perfectly applied in the case of *Usagi* (because the protagonists are two rabbits with short legs), it can be assumed that the size of the rabbits changes the feeling of the distance between them and their readers and that this leads to a change in the readers' viewpoint and their involvement in the fictive world of the picture book. For instance, the repetition of large and small sizes of the protagonists from Scenes 1 to 8 places the readers repeatedly at the locations of "far social" and "close social", which gradually leads them into the fictive world of the two rabbits. The close-ups of each rabbit in Scenes 9 and 10 put the readers at "far personal" distance, thus involving them more deeply in this fictive world, and the diminishing sizes of the protagonists from Scenes 11 to 13 gradually change the readers' position from "close social" to "far social" and even to "public", which would detach the readers from the world of the two rabbits.

### **3.1.3 Facial expression and eye direction of the characters**

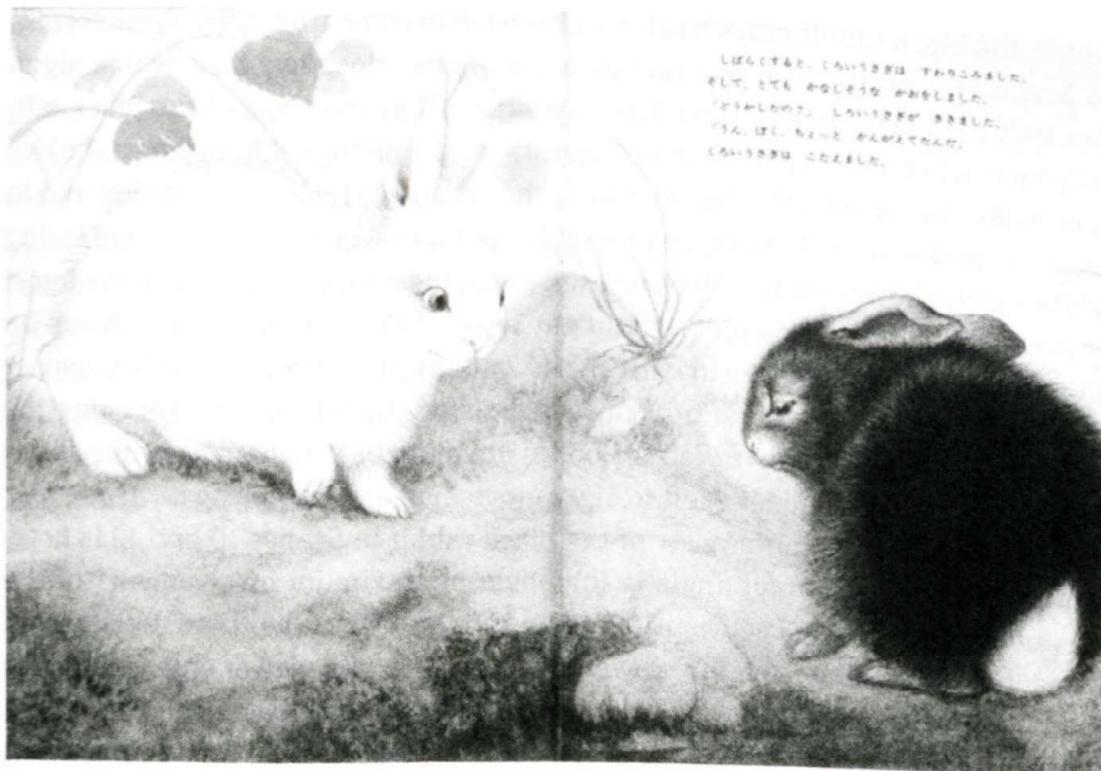
On all pages of the picture book *Usagi*, the characters' facial expressions (and their behaviours) are illustrated in realistic detail (see Figure 2). Since we have the ability to recognize and produce current emotions through facial expressions, such detailed facial expressions of the two rabbits can draw the readers' immediate attention and allow them to imagine how each character feels.<sup>4</sup>

As for the gaze or face (or sometimes body) direction of the characters, Nikolajeva and Scott (2001: ch. 5) have argued that a right-looking picture of a character shows that the character is going into the next stage of the story.<sup>5</sup> In the picture book *Usagi*, the face (or body) direction of the characters (especially that of the black rabbit) is closely linked to the speed of the progression of the relationship between the two rabbits. For instance, in Scenes 1, 2, 3 and 7, where the black rabbit directs his eyes (Scenes 1, 2 and 7) or his body (Scene 3) to the right (i.e. the same direction of reading), the text shows that the black rabbit enjoys playing

---

<sup>4</sup> It is well known that even newborn babies prefer to look at face-like configurations over non face-like ones. Snowden et al. (2012: ch.10) report that infants of 12 days old can recognize and respond to certain facial expressions and that by the first year infants can use the facial expressions of others as a guide to how one should behave in various situations.

<sup>5</sup> Nikolajeva and Scott (2001: ch.5) also discuss the relationship between the left and the right pages in double-page spreads. They argue that the left page means "home" or "secure" while the right portrays "away" or "adventure."



**Figure 2:** Scene 4 in *Usagi* (reproduced with permission)

with the white rabbit.<sup>6</sup> In these scenes, therefore, we can say that their relationship gradually deepens. In contrast, in Scenes 4, 5, 6 and 8, where the black rabbit directs his face or body to the left (i.e. the opposite direction of reading), the black rabbit cannot tell his wish to the white rabbit.<sup>7</sup> This means that their relationship does not progress in these scenes. In Scenes 9 and 10, the rabbits are shown alone in each scene for the first and only time, and this signifies a dramatic change in their relationship. As the text shows, in Scene 9 (where only the white rabbit is depicted with not only her eyes but also her face and body to the right), the white rabbit, who for the first time heard the black rabbit's wish in Scene 8, responds to his wish encouragingly, and, in Scene 10 (where only the black rabbit is illustrated with not only his eyes but also his face and body to the

<sup>6</sup> Nikolajeva and Scott (2001: ch.5) and Johnston (2012) have pointed out that picture book readers normally “read pictures” from left to right in accordance with their western reading conventions and that picture book authors draw pictures while taking this tendency for granted.

<sup>7</sup> A discrepancy can be found between the eye (or face) direction and the body direction in Scenes 6 and 8. In Scene 6, the black rabbit turns to the right with his body to the left, and in Scene 8 he turns to the left with his body to the right. Which of the two directions is more directly linked to the progression of the relationship between the two rabbits will be the subject of future work.

front), the black rabbit prays harder for his wish to come true. After these scenes, in Scene 11 onwards, the two rabbits are depicted close together, which signifies that the black rabbit's wish has come true. The directions of the black rabbit's face in Scenes 11 to 13 seem to suggest that the relationship between the two rabbits has become stable, because in Scenes 11 and 13 he looks to the left and in Scene 12 to the front. The illustration in Scene 14, in which the rabbits are facing to the right, shows that they have started living together in harmony in the forest into which the readers cannot set foot any longer.<sup>8</sup> This is more clearly shown on the back cover by the image of two rabbits walking away from the readers deeper into the forest. *Usagi*, therefore, conveys the story by effectively employing the correlations between eye (or face or body) direction of the characters, the speed of the story, and the readers' left-to-right movement across the pages.

In addition, the frontal view of the black rabbit in Scenes 10 and 12 is noteworthy. Since "the frontal angle is the angle of maximum involvement" of the viewers (Kress and van Leeuwen 2006: 145) and the gaze of the character depicted "demands something from the viewer, demands that the viewer enter into some kind of imaginary relation with him or her" (Kress and van Leeuwen 2006: 118), the frontal view of the black rabbit in these scenes, especially in Scene 10 in which the black rabbit is illustrated at the biggest size, would make the readers feel as if they were involved in the world of this picture book as participants.

### 3.2 What the text represents

Matsuoka (1987: 76–90), as an author, translator, and oral performer of picture books, has argued that the text of picture books is easy to understand when it contains some repetitions at regular intervals and consists of expressions from an objective point of view. In this section, therefore, I analyse the text of *Usagi* with special reference to the following expressions: (i) repetitions of a particular word/phrase/clause/paragraph; (ii) direct speech or direct quotation, which "evokes the original speech situation and conveys, or claims to convey, the exact words of the original speaker in direct discourse" (Coulmas 1986a: 2), (iii) deictic expressions like *kuru* ('come'), which are generally considered to reflect the speaker's viewpoint, and (iv) expressions which describe the characters' behaviours and/or facial expressions objectively. Expressions (i–iii) can be assumed to affect or determine the viewpoint of the readers, and expression (iv) evokes the readers' sensorimotor experience.

---

<sup>8</sup> The two rabbits in Scene 14 are no longer anthropomorphic. This signifies that the two rabbits have returned to the animal kingdom which we are not allowed to enter easily.

### 3.2.1 Repetitions

The text of *Usagi* exhibits repetitions of a particular phrase or clause, or even a particular paragraph.<sup>9</sup> A paragraph similar to that shown in (1) below, for instance, appears four times in this picture book, in Scenes 4, 5, 6 and from 7 to 8, although some changes are made in the second, third, and fourth repetitions; for example, the first and second sentences are conjoined like “... suwarikonde, totomo kanashisoona kao o shimashita” ('squatted down and showed a sad face') in the second and third repetitions and “kanashisoona kao o shite, suwarikonde shimaishita” ('showed a sad face and squatted down') in the fourth repetition; the phrase “Dooka shita no?” is changed to “Doo shita no?” in the second, third, and fourth repetitions; and the Japanese quotative marker *to* is added to the utterance of the black rabbit only in the fourth repetition.

- (1) *Shibarakusuru to kuroi usagi wa suwarikomimashita.*  
 while.pass when black rabbit TOP squat.down.POL.PAST<sup>10</sup>  
 ‘After a while, the black rabbit squatted down.’  
*Soshite totomo kanashisoona kao o shimashita.*  
 and very sad.look.ADN face ACC do.POL.PAST  
 ‘And he showed a very sad face.’  
 “*Dooka shita no?*” *Shiroi usagi ga kikimashita.*  
 what happen.PAST SFP white rabbit NOM ask.POL.PAST  
 ““What’s happened?” asked the white rabbit.’  
*“Un boku, chotto kangaeteta n da.” Kuroi usagi*  
 mm I little think.STAT.PAST NMZ COP black rabbit  
*wa kotaemashita*  
 TOP answer.POL.PAST  
 ““Mm, I’m just thinking,” answered the black rabbit.’ [Scene 4]

The phrase “kanashisoona kao o suru” also appears in the final scene, Scene 14, although it changes into the negative form. The form of [quoted sentence – reporting clause] in a direct quote is repeatedly used throughout the text (see also discussion of direct speech in 3.2.2), and in the conversation in Scene 11, the black rabbit repeats what the white rabbit says about their commitment to a shared

<sup>9</sup> For more on repetitions in picture books, including *Usagi*, see also Kasanuki (2010).

<sup>10</sup> The following abbreviations are used in the glosses: ACC = accusative, AND = adnominal, COM = complementizer, COP = copula, GEN = genitive, GER = gerundive, NEG = negative, NMZ = nominalizer, NOM = nominative, PAST = past, POL = polite, SFP = sentence-final particle, STAT = stative, TOP = topic.

future, using different intonation (shown by the different orthographical cues, “?” and “.” or “!”) to indicate an increasingly firm commitment. In addition, the phrase “me o manmaruku shite” (‘with perfectly round eyes’) appears twice, once in Scene 9 and again in Scene 10 (for more details of this expression, see Section 3.2.4), and the word “yattekuru” (‘come’) is used once in Scene 12 and once again in Scene 13 (see also 3.2.3 below).

Matsuoka (1987: 84–90) emphasizes that repetitions help the readers frame or set up the world described in a given narrative. Since *Usagi* shows many repetitions from the beginning of the story, it is reasonable to assume that this book makes it easy for readers to frame and understand the world of the two rabbits.

### 3.2.2 Direct speech

As Coulmas (1986b) and Yamaguchi (2009) have shown, the speech and thought representation (hereinafter, STR) in Japanese is quite different from that in English. Unlike English, which has several grammatically distinguishable types such as direct, indirect, and free indirect speech/thought (cf. Vandelanotte 2009), Japanese has no specialized grammatical forms which automatically distinguish direct and indirect discourse (see, for example, Yamaguchi 2009), although “there is a great variety of means indicating speaker perspective: directional and respectful or humble verbs, other lexical and morphological honorific personal pronouns, deictic demonstratives” (Coulmas 1986b: 172). In addition, Yamaguchi (2009 and personal communication) argues that the Japanese STR system reflects the reporter’s psychological distance from the content of the quoted speech or even from the reported speaker rather than the difference in perspective or viewpoint. Yamaguchi (2009) has pointed out, in discussing the differences between the quotative markers *to* and *tte* in Japanese (both of which are categorized as complementizers), that while the quotative marker *to* shows a neutral attitude by the reporter toward the speech reported, *tte* indicates that the reporter is detaching himself from the quoted speech or has no empathy with the speaker. Thus, while most of the research on English STR focuses on the ways in which different viewpoints (especially the narrator’s and the character’s viewpoints) are expressed and mixed (see, for instance, Rubba 1996; Sanders and Redeker 1996; Dancygier 2008; Vandelanotte 2009), the research on Japanese STR should be conducted with respect to the reporter’s psychological distance from the reported speech or from the reported speaker.

In *Usagi*, all of the utterances by the two main characters (22 utterances in total) are represented with quotation marks, i.e., are direct quotes, as shown in (1) above. Of these 22 utterances 5 appear with the quotative marker *to* and 17 appear

with no quotative marker, neither *to* nor *tte*.<sup>11</sup> The fact that the quotative marker *to* is used shows that the narrator is reporting the conversations between the two rabbits as he hears them, taking and maintaining a psychologically neutral stance to the contents or even to the two rabbits. As for the direct quotes with no quotative marker in Japanese, it has not been discussed even in Coulmas (1986b) and Yamaguchi (2009), but Yamaguchi (personal communication) suggests that the form adds a rhythm to the story and lets the readers read it smoothly. It seems to me that the frequent use of direct speech with no quotative marker in *Usagi* serves to take the readers effortlessly into the world of the two rabbits. The direct quotes in *Usagi* thus not only place the readers in a position where the conversation between the two rabbits can be heard, but also allow them to adopt and maintain a psychologically neutral stance to the protagonists and their utterances.<sup>12</sup>

Such a position, however, does not match the position reflected in the pictures. In 3.1.2, I argued that the readers' position could be assumed to change in accordance with the change in the size of the characters depicted. Section 4 below presents an investigation of which position is predominant when reading this picture book.

The book *Usagi* employs the quotative marker *tte* only once, as seen in (2):

- (2) *Sorekara to iu mono, kuroi usagi wa moo kesshite  
after.that COM say thing black rabbit TOP no.longer never  
kanashisoona kao o shimasen deshita tte.*  
sad.look.ADN face ACC do.POL.NEG COP.POL.PAST COM  
'And then, the black rabbit never showed a sad face, I heard.' [Scene 14]

This is the final sentence of the story. Given that *tte* shows the narrator's psychological detachment or remoteness from the content of the quoted speech, it is assumed that this *tte* serves as a prompt to make the readers detach themselves from the world of the two rabbits and return to the real world.

<sup>11</sup> Direct speech with no quotative marker is quite common in Japanese picture books.

<sup>12</sup> The polite form *masu* is consistently used in the narrative part of *Usagi*, and all of the verbs in the part, except *tobi-koeru* ('jump over') in Scene 3, appear in the past tense form *ta*. These facts also indicate that the narrator consistently adopts a neutral, objective stance toward the two rabbits and even the fictive world of the picture book.

### 3.2.3 Deictic expressions

In *Usagi*, only two deictic expressions, the demonstrative pronoun *kono* ('this') and the compound verb *yattekuru*, which consists of two verbs *yaru* ('send') and *kuru* ('come') but conveys almost the same meaning as the single verb *kuru* ('come'), appear in its narrative. These expressions indicate that the narrator's viewpoint is located close to the two rabbits. As shown in (3) and (4), both expressions are used in Scenes 12 and 13, i.e. in the scenes after the climax. Given that the direct quotes place the readers in a position where they can hear the two rabbits' conversation (see 3.2.2) and that the illustrations in the climax have the power to involve the readers in the fictive world of *Usagi* as participants (see 3.1.2 and 3.1.3), it can be reasonably assumed that readers would place themselves in the position very close to the protagonists (especially the white rabbit) in the climax and keep their position close to the protagonists after the climax. The use of the two deictic expressions in Scenes 12 and 13 is, therefore, quite natural, despite the two rabbits in the illustrations being depicted at a "far social" (Scene 12) or "public" (Scene 13) distance.

- (3) *Nihiki no kono shiawasena yoosu o mini, hoka no chiisana usagi ga oozei yattekismashita*  
 two.animal GEN this happy.ADN state ACC see.to other GEN small.ADN rabbit NOM many send.come.POL.PAST  
 'Many little rabbits came out to see how happy these two rabbits were.'

[Scene 12, my underlining]

- (4) *Mori ni sumu hoka no doobutsutachi mo dansu o oozei yattekimashita.*  
 forest in live other GEN animals also dance ACC many send.come.POL.PAST  
 'Other animals living in the forest also came out to see the dance'

[Scene 13, my underlining]

### 3.2.4 Onomatopoeia and other expressions

The text of *Usagi* describes the world of the two rabbits objectively, i.e. it mainly consists of descriptions of the actions and facial expressions of the two rabbits. The following sentence from Scene 9, for instance, perfectly describes the state of the white rabbit depicted in the picture of this scene.

- (5) *Shiroi Usagi wa me o manmaruku shite jitto kangaemashita*  
 white rabbit TOP eyes ACC perfectly.round do.GER steadily  
*kangaemashita*  
 think.POL.PAST

'The white rabbit steadily thought (about what the black rabbit said to her) with her eyes perfectly round.'

[Scene 9]



Figure 3: Scene 9 in *Usagi* (reproduced with permission)

The expression "me o manmaruku shite" completely matches the picture in this scene. The Japanese expression "me o maruku suru" (which literally means 'make one's eyes round') is commonly used as an idiom to express great surprise, and the prefix *man-* ('perfectly') attached to *maruku* stresses the magnitude of surprise. Therefore, by connecting what is said in the text and the image in Scene 9 to their own experience, the readers can easily understand how surprised the white rabbit was at what the black rabbit said to her. As for the word *jitto* (which is an example of onomatopoeia, although most Japanese people might not recognize the word as such), it shows that the white rabbit is frozen with surprise, in harmony with the picture of the white rabbit with her left hand placed in front of her mouth. Combining the text in (5) with the picture thus helps readers to understand more deeply the feelings of the white rabbit.

Onomatopoeia also serves as a prompt to arouse the reader's sensorimotor experience. As Fukada (2008) has argued, Japanese people use a variety of onomatopoeia in accordance with their own sensorimotor experience from the early stages of language acquisition.<sup>13</sup> Although TRW, the original English version of *Usagi*, has no onomatopoeia in it, the Japanese version contains several cases,

13 For a detailed discussion of Japanese onomatopoeia, see Kita (1997) and Tamori & Schourup (1999).

as listed under (6). The word *pyon* represents the jump of a small animal like a rabbit, *pyon pyon* the repetition of this action, and *pyoon* the act of jumping even higher and/or longer. The word *guruguru* expresses the repeated action of going around, *jitto* means ‘steadily,’ and *sotto* implies ‘gently.’

- (6) a. *pyon pyon no pyoon* [Scene 3, twice, ‘no’ is used as a linker here.]
- b. *pyon pyon* [Scene 6, once]
- c. *guruguru* [Scene 6, once]
- d. *jitto* [Scene 9, once]
- e. *sotto* [Scene 11, once]

## 4 Tension between words and pictures and how their contrasts are adjusted

As explained in Section 3, the illustrations and the text in the picture book *Usagi* show some discrepancies between them, especially in the viewpoint from which the world of the two rabbits is represented: while the illustrations show that the point of view is changing from scene to scene, the text presents a rather stable, objective viewpoint toward the world of the two rabbits. How can the readers deal with such discrepancies? The experiment conducted by Kojima et al. (2013) and its results can provide some answers to this question.<sup>14</sup>

### 4.1 An overview of the experiment by Kojima et al. (2013)

Kojima et al. (2013) performed an experiment to examine the effect of the text of *Usagi* on the viewpoint and gaze of the readers and their comprehension. In all, 22 undergraduate and graduate students at Kyoto University participated in this study (9 males and 13 females,  $M=20.91$  years,  $SD=1.98$  years). The participants were seated in front of a computer screen with their heads resting on a chin rest so that their eye movements could be measured. Each participant first performed a practice trial, and then two experimental trials. Two different types of stimulus sets were prepared for the experimental trials, both of which were based on the reduced scanned images of the double-spread pages of *Usagi* from the title page

---

<sup>14</sup> For a more detailed description of the method, see Kojima et al. (2013). Kojima et al. (2013), however, only gave a brief discussion of the results of the experiment. This paper analyses the data in more detail from a cognitive point of view.

to Scene 14 (15 images in total). The stimulus set used in the first experimental trial was a series of images with all of the words removed from the scanned data (hereinafter, "no-text condition"), and in the second set the images included both the text and illustrations (hereinafter, "text condition"). In both experimental trials, the stimuli were presented on the computer screen in the same order as in the actual picture book.

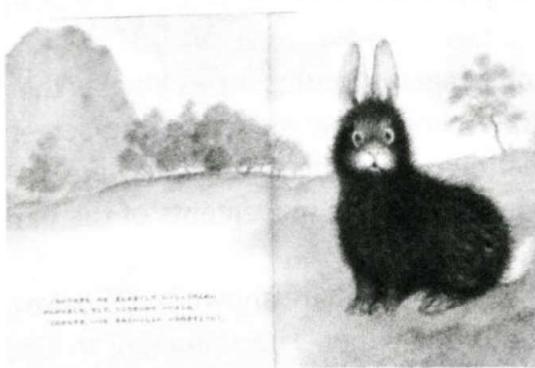
The participants were instructed to look at each stimulus for as long as they wanted and then to choose one of multiple choice statements on the screen describing their viewing experience and feelings. All of the statements were presented in Japanese. Throughout the experiment, the eye movements of the participants were recorded.

English translations of the statements we presented are shown in (7) below, although only those in (7b-d) are relevant to this paper. The statement in (7b) concerns the participants' sense of distance from the two rabbits, and the statements in (7c-d) involve the relative locations of the two rabbits.

- (7) English translations of the presented statements:
- [Title page] "You first looked at {the white rabbit/the black rabbit}".
  - [Scenes 1-8 and 11-13] "You feel as if {you are closer to the two rabbits/you are more distant from the two rabbits/you are staying at the same position}".
  - [Scene 9] "The black rabbit is {in front of/behind/to the left of/to the right of} the white rabbit".
  - [Scene 10] "The white rabbit is {in front of/behind/to the left of/to the right of} the black rabbit".
  - [Scene 14] "You feel that the two rabbits are {anthropomorphic/not anthropomorphic}".

We expected that in the no-text condition the participants' sense of distance would be affected only by the size of the characters (e.g., if the character is depicted as being larger than in the preceding scene, the participants would feel as if they were closer to the two rabbits), as discussed in 3.1.2; on the other hand, in the text condition we expected that the sense of distance would be greatly influenced by the text and that the participants would adopt a rather stable viewpoint, the same as that of the narrator (see discussion in 3.2.2). As for the relative positions between the two rabbits, the text shows that the white rabbit looking to the right in Scene 9 directs her eyes to the black rabbit (Figure 3) and that the black rabbit facing the front in Scene 10 looks at the white rabbit (Figure 4). If participants answered "front" in both scenes, it would indicate that the participants adopted a fictive position very close to the white rabbit, or even the same position

as the white rabbit's, and had deeper involvement in the two rabbits' world. We expected that in the text condition participants could correctly identify the position of each rabbit, while they could not in the no-text condition.



**Figure 4:** Scene 10 in *Usagi* (reproduced with permission)

## 4.2 The results of the experiment and discussion

### 4.2.1 The participants' sense of distance from the two rabbits

As Table 1 shows, statistically significant differences can only be observed in Scenes 1 and 6 between the no-text and the text conditions, but when focusing on the number of answers given as "the same", the number increases in Scenes 1, 3, 4, 5, 6, 8, 12 and 13 in the text condition. In particular, the increase in this answer and the decrease in the answer "more distant" in Scenes 3 and 5 deserve to be discussed, since the picture shows a viewpoint shift from "close social" to "far social" in these two scenes (see 3.1.2). In Scene 3, the text starts describing the two rabbits' conversation in the form of a direct quote, and in Scene 5, nearly the same conversation as in Scene 4 appears in the text (see 3.2.1). Assuming that being able to hear the conversation of others indicates physical proximity to them, the results in Scenes 3 and 5 would show that the direct quotations of the characters' utterances made some of the participants feel "not distant" from the two rabbits. The increase in "the same" and the decrease in "closer" responses in Scene 6 can also be explained in the same way. Since nearly the same conversation as in Scene 4 or 5 appears in this scene (see discussion in 3.2.1), the viewpoint of the text, i.e. the viewpoint fixed at the location at which the conversation between the two rabbits can be heard, overrides the viewpoint shift from "far social" to "close social" represented in the illustration. In Scenes 12 and 13, the increase in the number of "the same" responses in the text condition is slight and the

**Table 1:** Number of responses for each possible answer to the question about the participants' sense of distance

Condition	Scene	1	2	3	4	5	6	7	8	11	12	13
no-text	closer	15	16	8	14	5	18	2	15	12	9	5
	the same	3	3	3	3	2	4	6	2	3	2	2
	more distant	4	3	11	5	15	0	14	5	7	11	15
text	closer	4	18	9	14	4	9	3	18	13	7	5
	the same	6	3	8	4	8	8	4	4	2	4	4
	more distant	12	1	5	4	10	5	15	0	7	11	13
significance		*p < .01	ns	ns	ns	ns	*p < .01	ns	ns	ns	ns	ns

number of “more distant” is almost the same as that in the no-text condition. Since these scenes are the ones toward the end of the story, the participants who could understand the flow of the story from the text might feel as if they were gradually becoming distant from the world of the two rabbits, stimulated by the illustrations, even though the deictic expressions *kono* and *yattekuru*, both of which indicate that the narrator's viewpoint is close to the two rabbits (see 3.2.3), are used in these two scenes. The results show that when some conflict arises in the viewpoint between the text and the pictures, the text generally affects the readers' viewpoint more strongly than the pictures.

The results in Scene 1 under both conditions – 15 participants answered “closer” in the no-text condition and 12 answered “more distant” in the text condition – are surprising because the rabbits in this scene are depicted at almost the same size as those on the preceding title page. The results in the no-text condition might be attributed to the fact that the rabbits are depicted with detailed facial expressions (see discussion in 3.1.3). Since such facial expressions can be recognized at close range in everyday interaction, the participants might feel closer to the two rabbits despite the smallness of their size. In contrast, one plausible factor to explain the results in the text condition is that the participants recognized Scene 1 as the first page where the story begins. Since we know that the fictive world shown in picture books is not the same as the real world, the two rabbits, the citizens of the fictive world of *Usagi*, may have been regarded as distant.

#### 4.2.2 Relative locations of the two rabbits

Tables 2 and 3 show the responses to the questions about the relative locations of the two rabbits. The number of participants who said “front” in both scenes, Scenes 9 and 10, in the no-text condition was 8, but 17 in the text condition. The results indicate that the combination of text and pictures afforded a better understanding of these two scenes.<sup>15</sup> From Scene 1 to Scene 8, the pictures and the text, which portray matching and mismatching viewpoints, gradually lead the readers into the world of the two rabbits (see discussion in 3.1.2), and in the climax scenes, in Scenes 9 and 10, the text presents detailed descriptions of the protagonists which completely match the pictures depicted (see discussion in 3.2.4), and the combination between the words and the pictures enhances the readers’ involvement in the two rabbits’ world. The participants who read the text with pictures, therefore, were more easily involved in the world of the two rabbits in Scenes 9 and 10, as the illustrations induce them to do so (see discussion of the close-ups in 3.1.2 and the frontal view in 3.1.3), and correctly identified the relative locations of the two rabbits. Scenes 9 and 10 make good use of the dynamic interplay between words and illustrations.

**Table 2:** Number of responses for each possible answer to the question about the location of the black rabbit in Scene 9

Condition	front	behind	left	right
no-text	9	10	1	2
text	17	0	0	5

**Table 3:** Number of responses for each possible answer to the question about the location of the white rabbit in Scene 10

Condition	front	behind	left	right
no-text	15	3	1	3
text	21	0	1	0

<sup>15</sup> Before conducting the experiment, we had predicted that no participant would answer “behind” or “left” in these scenes; however, as the results show, some participants answered “behind” or “left” in the no-text condition in both scenes and one “left” even in the text condition in Scene 10. Although various plausible factors can be considered, a discussion of this is beyond the scope of the current paper.

## 5 Concluding remarks

Picture storybooks employ various strategies to prompt the readers to set up and understand the fictive worlds expressed within them. The current study has shown some of the strategies the picture book *Usagi* adopts, highlighting the dynamic interplay between the words, pictures and our everyday (social and physical) experiences. This study has also discussed in detail how the readers adjust the discrepancies between the words and the pictures in this picture book, based on the results of our experimental pilot study in 2013. However, there remain many issues to be solved: (i) undertaking a more detailed analysis of the experiment described in Section 4; (ii) conducting experiments on the effect of oral narratives on the (young) reader's viewpoint and understanding; (iii) exploring the differences in the readers' processing or understanding of the story based on the differences between the narrative styles of English and Japanese picture books; and (iv) modelling the process of understanding and/or engaging in the story. All of these issues will be addressed in future studies.

## References

- Coulmas, Florian. 1986a. Reported speech: Some general issues. In Florian Coulmas (ed.), *Direct and indirect speech*, 1–28. Berlin: Mouton de Gruyter.
- Coulmas, Florian. 1986b. Direct and indirect speech in Japanese. In Florian Coulmas (ed.), *Direct and indirect speech*, 161–178. Berlin: Mouton de Gruyter.
- Dancygier, Barbara. 2008. Personal pronouns, blending, and narrative viewpoint. In Andrea Tyler, Yiyoung Kim & Mari Takada (eds.), *Language in the context of use: Discourse and cognitive approaches to language*, 167–182. Berlin: Mouton de Gruyter.
- Dancygier, Barbara & Eve Sweetser (eds.). 2012. *Viewpoint in language: A multimodal perspective*. Cambridge: Cambridge University Press.
- Fauconnier, Gilles & Eve Sweetser (eds.). 1996. *Spaces, worlds, and grammar*. Chicago, IL: University of Chicago Press.
- Fukada, Chie. 2008. Embodiment and objectification in Japanese mimetics. In Kazuhiko Kodama & Tetsuharu Koyama (eds.), *Linguistic and cognitive mechanisms: Festschrift for Professor Masa-aki Yamanashi on the occasion of his sixtieth birthday*, 229–245. Tokyo: Hitsuj Shobo.
- Fukada, Chie. 2011. Ehon ni kakusareta shiza to shisen, shintaiteki keiken: *Shiroi Usagi to Kuroi Usagi no imisekai* [Viewpoint, gaze, and bodily experience in picture books: The world of *Shiroi Usagi to Kuroi Usagi*]. *Sapientia* [The St. Thomas University Review] 45. 157–174.
- Golden, Joanne M. & Annyce Gerber. 1990. A semiotic perspective of text: The picture story book event. *Journal of Reading Behavior* 22(3). 203–219.
- Johnston, Rosemary Ross. 2012. Graphic trinities: Languages, literature, and words-in-pictures in Shaun Tan's *The Arrival*. *Visual Communication* 11(4). 421–441.

- Kasanuki, Yoko. 2010. Ehon ni okeru ‘kurikaeshi’ to suki-ima [‘Repetition’ and Schema in Picture Books]. In Hiroshi Yoshioka, Kazuo Nakazawa, Shin’ichi Takeuchi, Shigeo Tonoike, Tomohiro Kawabata, Tadao Nomura & Shihoko Yamamoto (eds.), *Current studies for the next generation of English linguistics and philology: A Festschrift for Minoji Akimoto on the occasion of his retirement from Aoyama Gakuin University*, 157–169. Tokyo: Hitsuji Shobo.
- Kita, Sotaro. 1997. Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35. 379–415.
- Kojima, Takatsugu, Chie Fukada, Teppei Tanaka & Masashi Sugimoto. 2013. Ehon dokusha no shiza/shisen ni honbun ga oyobosu eikyo: Ehon *Shiroi Usagi to Kuroi Usagi* o mochiita gankyu-undo sokutei ni yoru kento [The effect of text on the viewpoint and gaze of a picturebook reader: From the measurement of eye movement of the readers of *Shiroi Usagi to Kuroi Usagi*]. *Proceedings of the Japanese Society for Cognitive Psychology: The 11th Conference of the Japanese Society for Cognitive Psychology* 64.
- Kress, Gunther & Theo van Leeuwen. 2006. *Reading images: The grammar of visual design*. London: Routledge.
- Matsuoka, Kyoko. 1965. *Shiroi Usagi to Kuroi Usagi* [The white rabbit and the black rabbit]. Tokyo: Fukuinkan Shoten.
- Matsuoka, Kyoko. 1987. *Ehon no sekai, kodomo no sekai* [The world of picture books, the world of children]. Tokyo: Japan Editors School Press.
- Nakamura, Toshie. 1995. Ehon no yomikikase ni kansuru shinrigakuteki kenkyu (III): Yoji no monogatari rikai ni oyobosu wadaijoho no eikyo [Psychological studies on reading picture books to young children (III): Effects of topic information on young children’s story comprehension]. *The Science of Reading* 39(1). 16–23.
- Nikolajeva, Maria & Carole Scott. 2001. *How picturebooks work*. New York: Routledge.
- Oka, Natsuki, Xia Wu, Chie Fukada & Motoyuki Ozeki. 2013. Concurrent acquisition of the meaning of sentence-final particles and nouns through human-robot interaction. In Minho Lee, Akira Hirose, Zeng-Guang Hou & Rhee Man Kil (eds.), *Neural information processing: 20th International Conference, ICONIP 2013, Daegu, Korea, November 3–7, 2013, Proceedings, Part I (Lecture Notes in Computer Science)*, 387–394. Heidelberg, Germany: Springer.
- Rubba, Jo. 1996. Alternate grounds in the interpretation of deictic expressions. In Gilles Fauconnier & Eve Sweetser (eds.), *Spaces, worlds, and grammar*, 227–261. Chicago, IL: University of Chicago Press.
- Sanders, José & Gisela Redeker. 1996. Perspective and the representation of speech and thought in narrative discourse. In Gilles Fauconnier & Eve Sweetser (eds.), *Spaces, worlds, and grammar*, 290–317. Chicago, IL: University of Chicago Press.
- Snowden, Robert, Peter Thompson & Tom Troscianko. 2012. *Basic vision: An introduction to visual perception*. Oxford: Oxford University Press.
- Tamori, Ikuhiro & Lawrence Schourup. 1999. *Onomatope: Keitai to imi* [Onomatopoeia: Form and meaning]. Tokyo: Kuroshio Shuppan.
- Vandelanotte, Lieven. 2009. *Speech and thought representation in English: A cognitive-functional approach*. Berlin: Mouton de Gruyter.
- Williams, Garth. 1958. *The Rabbits’ Wedding*. New York: HarperCollins Publishers.
- Yamaguchi, Haruhiko. 2009. *Meisekina inyo, shinayakana inyo: Waho no nichi-ei taisho kenkyu* [Quotational clarity and flexibility: A contrastive study of speech and thought representation in English and Japanese]. Tokyo: Kuroshio Shuppan.