MULTIMODAL TEXT-IMAGE SYNERGY
IN REPRESENTING INTERPERSONAL RELATIONS IN PICTURE BOOKS

Victoria Yefymenko
Doctor of Sciences in Linguistics, Associate Professor,
Taras Shevchenko National University of Kyiv
(60, Volodymyrska St., Kyiv, 01033, Ukraine);
email: v.yefymenko@knu.ua
ORCID: https://orcid.org/0000-0001-8135-4983

Abstract
Picture books as multimodal narratives contain rich data for analysis of multimodal meaning-making. Verbal and visual modes offer various choices for representing different interpersonal relations. In this paper, I focus on the following research question: how visual and verbal resources are used to construct interactive meanings between the reader and characters and between the characters of picture books.
In my analysis, I follow Halliday's social semiotic approach to language within systemic-functional linguistics, according to which every semiotic system operates at three levels corresponding to ideational, interpersonal and textual metafunctions. I focus on the interpersonal level, using methodology and typologies from Kress and van Leeuwen's (2006) visual grammar, further developed by Painter, Martin, and Unsworth (2012) for the analysis of picture books. Visual resources used for representing interpersonal relations are connected with the categories of contact, social distance, attitude, pathos, affect, ambience, and graduation. Verbal means of realizing interpersonal relations include, among others, mood structures and attitudinal lexis. Martin and White's (2005) appraisal theory is applied to classify appraisal resources according to three systems: attitude, engagement and graduation.
I have analysed interpersonal text-image relationship in two picture books: *The Paper Bag Princess* by Munsch (2005) and *Snow White in New York* by French (1989). The analysis of visual resources at the interpersonal level has demonstrated a predominance of long shots, direct eye contact with a reader is rare with the exception of several close-ups. Front-on images situate characters in an equal position with readers, creating a sense of involvement. Overall, the interactions between the visual and verbal components create a synergetic effect in representing interpersonal relations in picture books.
Keywords: *multimodality, text-image relationship, picture book, visual grammar, interpersonal metafunction.*

1. Introduction
Multimodal studies are a rapidly developing field of research focusing on the role and contribution of various modes in the meaning making process (O'Halloran & Lim, 2014; Ledin & Manchin, 2020). Special attention is paid to the interaction between verbal and visual modes (Martinec & Salway, 2005; Bateman, 2014; Otto et al., 2020). At the same time researchers of films (Tseng, 2023) add an aural mode to their analysis, and van Leeuwen (2020, p. 498) believes that "recently typography has become a communicative mode in its own right". A change of focus from language to other modes,

in particular, a visual mode, has resulted in seminal publications on illustrations in picture books (Nikolajeva & Scott, 2001; Lewis, 2001; Painter, 2017) and the grammar of visual design (Kress & van Leeuwen, 2006). These seminal works take as a starting point Halliday's systemic functional linguistics. Halliday (1978) identifies three language metafunctions: ideational, interpersonal, and textual. The ideational metafunction refers to the linguistic representation of action (who is doing what, to whom, when, and where). The interpersonal metafunction describes interactions between participants (such as asking questions, making statements, or giving commands) at the linguistic level. And the textual metafunction describes the flow of information within and between texts, including how texts are organized, what is made explicit and what is assumed as background knowledge, how the known and the new are related, and how coherence and cohesion are achieved.

In this study, I will focus on the interpersonal metafunction. Kress and van Leeuwen (2006, p. 42) believe that for multimodal narratives the interpersonal metafunction refers to the relations between the producer of a sign, the viewer, and the object represented. But I will follow Painter, Martin, and Unsworth (2012) in their broader interpretation of the interpersonal metafunction as such that expresses affiliation and feelings both between the reader/ viewer and character and between characters in the story.

The *aim* of this present study is to analyse the interpersonal relations in picture books, created by the interplay between verbal and visual modes. In accordance with this aim the following *tasks* should be completed: to outline the methodological approaches to the analysis of interpersonal relations in picture books; to identify elements of the appraisal system in English, which are widely used in the representation of the interpersonal relations; and to specify how various visual categories interrelate with verbal expressions of mood and appraisal. *The subject-matter* of our study is visual and verbal meaning-making tools at the interpersonal level. *The material* for the analysis has been taken from two picture books: *Snow White in New York* by Fiona French (1989) and *The Paper Bag Princess* by Robert Munsch (2005).

Though I admit that there might be individual cases of discrepancy between meanings expressed visually and verbally, my hypothesis is that the interactions between the visual and verbal components create a synergetic effect in representing interpersonal relations in picture books.

2. Method

I will examine the visual and verbal components of interpersonal meaning. Based on Kress and van Leeuwen (2006) and Painter, Martin, and Unsworth (2012) typologies, I will focus on the following visual categories: contact, social distance, attitude, pathos, affect, vibrancy, warmth, familiarity, and graduation. These categories will be discussed in more detail in the Results section.

Interpersonal choices in the verbiage focus mainly on mood structures and attitudinal lexis (Santamaría-García, 2022, p. 149). Mood structures, as is well known, include indicative mood (namely, interrogative and declarative sentences) and imperative mood. Connected with mood are speech acts. According to Halliday (1985), speech acts (or speech functions as he calls them) can (1) offer information, which are realized by statements in the indicative mood; they can (2) offer goods-and-services, expressed by questions or commands (e.g., Would you like a drink?); they can (3) demand information, i.e., form a question, which is realized by the interrogative mood; and they can (4) demand goods-and-services, expressed by commands. Kress and van Leeuwen applied Halliday's theory to images. "When images 'offer', they primarily offer information. … When images 'demand', they demand, one could say, the 'goods-and-services' that realize a particular social relation." (Kress & van Leeuwen, 2006, p. 123).

As far as interpersonal meaning is concerned, a verbal mode is closely connected with the expression of appraisal. I will refer to the appraisal theory (Martin & White, 2005). The authors of the appraisal theory identify three major systems: the attitude system, the engagement system, and the graduation system. The attitude system includes such categories as appreciation, judgement, and affect. By appreciation Martin and White mean assessment of artifacts, entities, happenings, and states

of affairs by reference to aesthetics and other systems of social valuation. Judgment refers to assessment of human behavior and/or character by reference to ethics and other social norms. And, finally, affect treats assessment as an emotional reaction (White, 2015, p. 3). The engagement system covers instances, when the authorial voice is 'engaged' with the other voices and alternative positions in the current communicative context. And, finally, locutions, grouped together under the heading of 'graduation', are used by writers to 'graduate' either the force of the utterance or the focus of the categorisation by which semantic values are identified.

3. Results

The first category to be analysed is *contact*. It has already been mentioned that according to Kress and van Leeuwen (2006) pictures can either offer information or demand some action from the reader. In *The Paper Bag Princess* there are only 'offer' pictures, where participants do not look directly at the viewer, and in *Snow White in New York* there are several 'demand' pictures (which is a visual form of direct address). For instance, a picture of Snow White's father and stepmother's wedding contains both demand (stepmother) and offer (father) images. They are accompanied by an offer speech act, expressed by a statement in the indicative mood:

(1) "*But one day he married again.*" (French, 1989, p. 4).

Another example of demand is a medium shot of the stepmother, accompanied by a text:

(2) "*And she plotted to get rid of her stepdaughter.*" (French, 1989, p. 9).

It is again, as in the previous example, an 'offer information' speech act. There is a mirror image of Snow White on the other part of the double page spread, not looking at the reader, which makes it an 'offer' picture'. It is accompanied by a text, offering the following information:

(3) "*Snow White the Belle of New York City.*" (French, 1989, p. 8).

A conversation between Elizabeth and the dragon from *The Paper Bag Princess* is accompanied by 'offer' pictures and contains a number of interrogative and imperative sentences. First, the dragon tries to persuade the princess to go home and come back the next day, when he is hungry again. Such speech acts as "*go away*", "*come back*" demand goods-and-services and are the indication of power. But the girl manages to outwit him by asking tricky questions demanding information:

(4) "*Is it true that you are the smartest and fiercest dragon in the whole world?*" (Munch, 2005, p. 14);
(5) "*Is it true that you can burn up ten forests with your fiery breath?*" (Munsch, 2005, p. 16);
(6) "*Dragon, is it true that you can fly around the world in just ten seconds?*" (Munsch, 2005, p. 20).

When a girl starts demanding services by using imperative sentences "*do it again!*" (Munsch, 2005, p. 20), she appears in the position of power.

Another category, named by Kress and van Leeuwen, is *social distance*, realised by a shot size. In *The Paper Bag Princess,* all shots are long, characters are presented as full-length images, while in *Snow White in New York* we find a close-up, medium shots and long shots. The close-up depicts the stepmother, when she has decided to get rid of Snow White herself. There is a verbal expression of attitude: affect: dissatisfaction "*was mad with rage*" (French, 1989, p. 19) (here and hereafter Martin & White's [2005] classification is used). Usually "a sense of intimacy between viewer and character is created" (Painter et al., 2012, p. 16) in close-up shots, but it is not the case.

One more category is *attitude*, which consists of involvement and power. The degree of involvement depends on the angle from which a character is viewed. In *The Paper Bag Princess,* characters are mainly presented facing the reader 'sideways on', while in *Snow White in New York,* there are several 'front on' images, which create "a maximum sense of involvement" (Painter et al., 2012, p. 17), e.g., Snow White singing on stage. In the preceding dialogue between Snow White and seven jazz-men interrogative and imperative sentences are used, which offer goods-and-services ("*Stay here if you like ...*" (French, 1989, p. 14), demand information ("*What can I do?*"; "*Can you sing?*" (French, 1989, p. 15), and demand goods-and-services ("*you'll have to work*" (French, 1989, p. 14). They are used to offer help and clarify in what way this help can be delivered, thus creating interactive relationships between characters. Jazz-men's attitude is also expressed: attitude: affect: unhappiness ("*were sorry for her*" (French, 1989, p. 14). The same attitude is expressed in one more 'front on' image of Snow White, but this time lying in the coffin, carried by jazz-men. The accompanying text says:

(7) "*The seven jazz-men, their hearts broken, carried the coffin unsteadily up the church steps.*" (French, 1989, p. 28).

The category of *affect* includes emotions depicted in facial expressions, gestures, and bodily stance of characters. Comparing characters' emotions in picture books and manga, I can point out that manga are characterized by a greater distortion of faces as a result of extreme emotional states of characters (Yefymenko, 2022). The first image from *The Paper Bag Princess* depicts princess Elizabeth and prince Ronald. They are drawn close to each other, but Ronald has turned his back to Elizabeth. So, the orientation (and his facial expression) indicate that they are not as close as the girl thinks. The accompanying text contains appraisal, namely attitude: appreciation: reaction: positive and engagement: expand: entertain – an example of the epistemic modality:

(8) "*She lived in a castle and had expensive princess clothes. She was going to marry a prince named Ronald.*" (Munsch, 2005, p. 6).

Elizabeth and the dragon are also depicted rather close to each other in almost all the images, but it does not signify their closeness. I have already analysed their conversation in terms of a category of contact. Applying the appraisal theory, I can find here examples of the attitude system, namely judgement: social esteem (*fantastic*, *magnificent*), and the graduation system: intensification: quality (*the smartest and fiercest*). All these adjectives are used by Elizabeth to trick the dragon.

The second dialogue between Elizabeth and Ronald reflects very clearly a superior attitude of the prince and bewilderment of Elizabeth. It should be mentioned that *The Paper Bag Princess* uses a minimalist drawing style, but even such images can depict emotions in facial expressions and bodily stance. Ronald tries to demonstrate his power over Elizabeth by demanding services with the help of imperative sentences:

(9) "*Elizabeth, you are a mess! You smell like ashes, your hair is all tangled and you are wearing a dirty old paper bag. Come back when you are dressed like a real princess.*" (Munsch, 2005, p. 26).

But he fails because the girl has the last word. A verbal mode expresses attitude: appreciation: reaction: negative. Interestingly, the last phrase in this book, told by Elizabeth, contains a positive reaction, as well as graduation: intensification: degree (*really pretty*, *very neat*), but she finally understands the true nature of Robert and is not going to marry him:

(10) *"Ronald," said Elizabeth, "your clothes are really pretty and your hair is very neat. You look like a real prince, but you are a bum."* (Munsch, 2005, p. 28).

This picture book provides an alternative to classic fairy tales "by filling the gap with an active heroine who chooses not to conform to stereotypical gender patterns" (Joosen, 2011, p. 47).

The significance of *colour* in expression of interpersonal relations is really high, colours produce an emotional response in viewers. Scholars (Kress & van Leeuwen, 2002; Kress & van Leeuwen, 2006; Painter et al., 2012) analyse colour in terms of modality and other categories, such as ambience, which includes vibrancy, warmth, and familiarity.

Starting with *vibrancy* (or the depth of saturation), one can notice high saturation of images in *Snow White in New York* and less saturated colours of *The Paper Bag Princess*. One of the double spread pages in *Snow White in New York* is almost completely dark, depicting the funeral of Snow White and creating the corresponding emotional reaction of the reader – sadness. And at the linguistic level there is an expression of attitude: affect: unhappiness:

(11) *"All New York was shocked."* (French, 1989, p. 25).

The dark ambiance of the funeral is in a sharp contrast with the warm ambiance of a wedding party.

In the second ambience system – that of *warmth* – the warmer colours are in contrast with cooler colours. In *Snow White in New York,* there are images depicted almost entirely in warm colours and in cool blue colours. For instance, the performance of jazz-men is presented in yellow and red colours, while the reporter, who saw Snow White singing for the first time, is drawn in different shades of blue colour. The reporter is described in cool colours when he is a stranger, but later instead of a blue suit he would wear a brown one – a warm colour, indicating his warm feelings to Snow White. Besides, we come across relations of engagement, expressing epistemic modality: expand: entertain:

(12) *"He knew at once that she would be a star."* (French, 1989, p. 16).

One more feature worth mentioning is "the use of contrastive ambient splashes" (Painter et al., 2012, p.40), e.g., a pink dress of Snow White and flowers in the coffin provide a light splash in the overall dark ambiance of the funeral scenes.

Such category as *graduation* is of a special nature, because it can be applied to the analysis of both visual and verbal means. Painter, Martin, and Unsworth (2012) borrowed this term from Martin and White's appraisal theory. It refers "to strategies for intensifying gradable interpersonal meanings" (Painter et al., 2012, p. 18) and includes scaling and quantification. Visually it is expressed, for instance, by depicting a large number of people on the street, and a phrase "*crowds of people stood in the rain*" (French, 1989, p. 27) is an example of verbal graduation: force: quantification. Apart from quantification such verbal expression of graduation as intensification may be used. For instance, the stepmother is called "*the classiest dame in New York*" (French, 1989, p. 6), which is the intensification of quality.

4. Discussion
Comparing the category of affect, expressed by visual and verbal means, it has been found out that in *Snow White in New York* there are instances, when affect is expressed only verbally. It happens either when the character, whose emotions are described, is represented as a very small image (e.g., a tiny figure of Snow White against the background of the skyscrapers of New York), where no facial expression is visible, or as silhouettes (e.g., jazz-men in the jazz club or crowds of people mourning Snow White's death). At the same time there are instances, when affect is expressed by both visual and verbal devices. Namely, there are clear indications of rage in a close-up of the stepmother, which are amplified by the text:

(13) "*The stepmother was mad with rage.*" (French, 1989, p. 19).

In another example from the same picture book, there is a double spread page with mirror images of seven jazz-men carrying the coffin with Snow White. In the first image, sadness is reflected in their facial expressions and further stressed in the text:

(14) "*The seven jazz-men, their hearts broken, carried the coffin unsteadily up the church steps.*" (French, 1989, p. 28).

In the second image their amazement and joy is expressed visually (all of them are smiling), as well as verbally:

(15) "*Suddenly one of them stumbled, and, to everyone's amazement, Snow White opened her eyes.*" (French, 1989, p. 28).

Such category as pathos, which is determined by a drawing style, is minimalist in *The Paper Bag Princess* and generic in *Snow White in New York* with more expressive eyes and mouth indicating a wider range of emotions. A generic style is considered to be more empathic, but our analysis gives evidence that both types are engaging. For instance, a reader empathises with Elizabeth in her confrontation with the dragon, though there are no close-up and front-on shots of her.

Certain categories and connected with them issues require further discussion. A rather controversial question connected with categories of social distance and attitude is a degree of intimacy and involvement, experienced by a reader. Some formal criteria, such as close-up, front-on shots, which presuppose a maximum degree of intimacy and involvement, do not always work in practice, for instance, if they are connected with the depiction of antagonists (as it happens in the case of Snow White's stepmother).

5. Conclusions

This study has revealed that there is no strict interdependence between 'offer' and 'demand' pictures and speech acts: 'demand' pictures may be accompanied by 'offer information' speech acts, while 'offer' pictures may come in pairs with 'demand goods and services' speech acts. Another conclusion is that a sense of intimacy between a viewer and a character is not always created in close-up shots.

The results of my analysis have demonstrated the complementarity and interdependency of visual and verbal meaning in the construction of interpersonal relations between the reader/viewer and character and between characters in the story. Text and images have different resources to create engagement and different attitudinal meaning, but the combined multimodal meaning resulting from their interaction creates, as a rule, a synergetic effect.

Issues for ***future possible research*** may include the analysis of interactions between visual and verbal modes at the ideational and textual levels, as well as relations between verbal, visual, and aural modes in digital picture books.

Declaration of competing interest
The author has no conflicts of interest to declare that are relevant to the content of this article.

References
Bateman, J. (2014). *Text and image: A critical introduction to the visual/verbal divide*. London: Routledge.
Halliday, M.A.K. (1985). *An introduction to functional grammar*. London: Edward Arnold.
Halliday, M.A.K. (1978). *Language as social semiotic*. London: Edward Arnold.

Joosen, V. (2011). *Critical and creative perspectives on fairy tales: An intertextual dialogue between fairy-tale scholarship and postmodern retellings*. Detroit, MI: Wayne State University Press.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. London: Routledge.

Kress, G., & van Leeuwen, T. (2002). Colour as a semiotic mode: Notes for a grammar of colour. *Visual communication, 1*(3), 343–369.

Ledin, P., & Machin, D. (2020). *Introduction to multimodal analysis*. London: Bloomsbury.

Lewis, D. (2001). *Reading contemporary picture books: Picturing text*. London: Routledge.

Martin, J.R., & White, P.R.R. (2005). *The language of evaluation: Appraisal in English*. Basingstoke & New York, NY: Palgrave MacMillan.

Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual communication, 4*(3), 337–371.

Nikolajeva, M., & Scott, C. (2001). *How picture books work*. New York & London: Garland Publishing.

O'Halloran, K.L., & Lim, F.V. (2014). Systemic functional multimodal discourse analysis. In S.Norris & C.D.Maier (Eds.), *Interactions, images and texts: A reader in multimodality* (pp. 137–154). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781614511175.137

Otto, C., Springstein, M., Anand, A., & Ewerth, R. (2020). Characterization and classification of semantic image-text relations. *International Journal of Multimedia Information Retrieval*, *9,* 31–45. https://doi.org/10.1007/s13735-019-00187-6

Painter, C. (2017). Multimodal analysis of picture books. In B. Kümmerling-Meibauer (Ed.), *The Routledge companion to picture books* (pp. 420–428). London: Routledge. https://doi.org/10.4324/9781315722986

Painter, C., Martin, J.R., & Unsworth, L. (2012). *Reading visual narratives: Image analysis in children's picture books*. Sheffield: Equinox.

Santamaría-García, C. (2022). A semiotic and multimodal analysis of interactive relations in picture books that challenge female gender stereotypes. In A. J. Moya-Guijarro & E. Ventola (Eds.), *A multimodal approach to challenging gender stereotypes in children's picture books* (pp. 144–163). New York & London: Routledge. https://doi.org/10.4324/9781003145875

Tseng, C. (2013). *Cohesion in film: tracking film elements*. Basingstoke: Palgrave Macmillan.

Van Leeuwen, T. (2020). Multimodality and multimodal research. In L. Pauwels & D.Mannay (Eds.), *The SAGE handbook of visual research methods* (pp. 465–508). Thousand Oaks, CA: SAGE publications, Inc. https://doi.org/10.4135/9781526417015

White, P.R.R. (2015). Appraisal theory. In K. Tracy, C. Ilie, & T. Sandel (Eds), *The international encyclopedia of language and social interaction*. Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781118611463.wbielsi041

Yefymenko, V. (2022). Multimodality and cross-modal cohesion in manga. *Cognition, communication, discourse*, *24*, 103–114. https://doi.org/10.26565/2218-2926-2022-24-08

Sources for illustrations

French, F. (1989). *Snow White in New York*. Oxford: Oxford University Press.

Munsch, R. (2005). *The Paper Bag Princess*. New York, NY: Annick Press Ltd.

МУЛЬТИМОДАЛЬНА СИНЕРГІЯ ТЕКСТУ ТА ЗОБРАЖЕННЯ У РЕПРЕЗЕНТАЦІЇ ІНТЕРПЕРСОНАЛЬНИХ ВІДНОСИН В ІЛЮСТРОВАНИХ КНИЖКАХ

Вікторія Єфименко
доктор філологічних наук, доцент,
Київський національний університет імені Тараса Шевченка
(60, Володимирська вул., м. Київ, 01033, Україна);
email: v.yefymenko@knu.ua
ORCID: https://orcid.org/0000-0001-8135-4983

Анотація
Ілюстровані дитячі книжки як мультимодальні наративи містять багатий матеріал для аналізу способів мультимодального смислотворення. Вербальний і візуальний модуси пропонують власні засоби для репрезентації різних видів інтерперсональних відносин. Ця стаття вирішує таке дослідницьке питання: як візуальні та вербальні ресурси використовуються для побудови інтерактивної взаємодії між читачем і героями творів, а також між самими персонажами ілюстрованих книжок.

У цьому аналізі використано соціосеміотичний підхід Хеллідея до мови в рамках системно-функціональної лінгвістики, відповідно до якого кожна семіотична система функціонує на трьох рівнях, що відповідають ідеаційній, інтерперсональній і текстовій метафункціям. Стаття зосереджена на інтерперсональному рівні, використовуючи методологію та класифікації візуальної граматики Кресса та ван Левена (2006), яку Пейнтер, Мартін та Ансворт (2012) адаптували для аналізу ілюстрованих дитячих книжок. Візуальні ресурси, що використовуються для репрезентації інтерперсональних відносин, пов'язані з такими категоріями, як контакт, соціальна дистанція, ставлення, пафос, афект, атмосфера й градація. Вербальні засоби реалізації інтерперсональних відносин включають, серед іншого, засоби вираження модальності та ставлення. У зв'язку з цим застосовано теорію оцінки Мартіна та Уайта (2005) для класифікації оцінювальної лексики відповідно до трьох систем: ставлення, залученість і градація.

У роботі проаналізовано інтерперсональні взаємовідношення між текстом і зображенням у двох дитячих ілюстрованих книжках: «Принцеса в паперовому пакеті» Мунша (2005) і «Білосніжка в Нью-Йорку» Френч (1989). Аналіз візуальних ресурсів на інтерперсональному рівні продемонстрував переважання зображень із загальним планом, прямий зоровий контакт із читачем зустрічається рідко, за винятком кількох крупних планів. Зображення в анфас створюють ефект рівності персонажів з читачами та відчуття причетності. У цілому, взаємодія між візуальними та вербальними компонентами створює синергетичний ефект у побудові інтерперсональних відношень в ілюстрованих дитячих книжках.

Ключові слова: *мультимодальність, взаємовідношення між текстом і зображенням, ілюстрована дитяча книжка, візуальна граматика, інтерперсональна метафункція.*