NURSERY RHYMES: A COMPARATIVE STUDY FROM SYSTEMIC FUNCTIONAL-MULTIMODAL DISCOURSE ANALYSIS PERSPECTIVE

Ton Nu My Nhat[*]

Quy Nhon University

Abstract

The contemporary landscape of communication is marked by the ubiquity of multimodality. However, the study of multimodal discourse is a relatively recent domain of enquiry. This paper is aimed to contribute to the current exploration of this emerging field. This paper is situated within the Systemic functional-Multimodal discourse analysis (SF-MDA) approach. The first part will explicate the aspects of Systemic Functional Theory which are pertinent to discourse of multimodal nature. Drawing on recent works, the second part will present a brief account of the SF-MDA approach. The metafunctional systems of image analysis of children's picture books will also be presented, serving as a framework for the comparative analysis of books of nursery rhymes (NR) in English and Vietnamese to be described. The paper concludes with a brief discussion of the pedagogical implications of SF-MDA in the Vietnamese educational context.

Key words: systemic functional theory, discourse analysis, multimodal discourse analysis

1. Introduction

With the rapid technological advances, the landscape of inscribed communication has accordingly changed, with the co-deployment of both images and language as closely-intertwined meaning-making resources. It is now widely accepted that discourse studies can no longer be confined to the realm of language alone. However, the study of multimodal discourse is very much in infancy the world over, and understandably is still under-researched in the Vietnamese context. This paper is, therefore, aimed to contribute to the current exploration of this emerging field. The subject of the study is the picture books of nursery rhymes in English and Vietnamese. The decision to choose this genre lies with its role as a significant instrument of socialization for pre-school children. C. Painter, J. Martin & L. Unsworth (2013) state that children's picture books can be recognized as "*a key means of apprenticeship into literacy, literature and social values, which in turn means that how they are constructed to accomplish these ends is an important educational question. Since most of the space in picture books is given over to the pictures, it is reasonable to suppose*

---

[*] Email: tnmynhat70@gmail.com

that the visual component of the books, quite as much as the verbal, is crucial in this apprenticeship." (p.1)

This presentation is situated within the Systemic Functional (SF) – Multimodal Discourse Analysis (MDA) approach. The paper will start with the aspects of Systemic Functional Theory (SFT) which are pertinent to discourse of multimodal nature. Drawing on recent works, the second part will present a brief account of the SF-MDA approach. The metafunctional descriptive systems of image analysis of children's picture books will also be summarized, serving as a framework for the comparative analysis of books of nursery rhymes in English and Vietnamese to be described. The paper concludes with some pedagogical implications of SF-MDA in the Vietnamese educational context.

2. Theoretical background

2.1. SF model for language

The idea of multimodality began as early as the 4th century BC, when classical rhetoricians referred to it with their emphasis on voice, gesture, and expressions in public speaking. However, it was not until the mid 1990s, with the publications of O'Toole's (1994/2010) *Language of Displayed Art* and Kress & van Leeuwen's (1996/2006) *Reading Images: the Grammar of Visual Design* did the term begin to gain much attention. From then onwards, there has been a remarkable increase in interest in the analysis of modes other than language. There are different approaches to MDA, as comprehensively reviewed in seminal works on MDA. As this paper is situated in SF-MDA, the following sections will be concerned with only aspects relevant to this approach.

There are several reasons why the Hallidayan model of Systemic Functional Linguistics (hereafter SFL) has proven itself useful as a theoretical framework for the analysis of intersemiotic relations in multimodal texts.

Table 1. SF framework for (a) language (Halliday, 2004; Martin, 1992) and (b) visual images (based on O'Toole, 1994)

	(a) LANGUAGE (b) VISUAL IMAGES	
CONTENT Stratum	**Discourse Semantics** Discourse Relations (paragraph and text) **Lexicogrammar** Clause complex Clause Word group Word (Metafunctionally Based Systems)	**Discourse Semantics** Intervisual Relations Work **Grammar** Scene Episode Figure Part (Metafunctionally Based Systems)
EXPRESSION Stratum	**Typo-/Graphology and Phonology** (Cross-Functional Systems)	**Graphics** (Cross-Functional Systems)

(Source: O'Halloran, 2008, pp. 449-451)

Central to the SFT is making, understanding and evaluating meanings as they are used in context. Halliday (1994, p. xiii) explicates that the use of the term 'functional' in SFT is "because the conceptual framework on which it is based is a functional one rather than a formal one". According to SFL, the fundamental purpose that language has evolved is to enable us to make meanings. SFL proposes that these meaning-making functions can be grouped into three main categories, or metafunctions – the IDEATIONAL, the INTERPERSONAL, and the TEXTUAL. These three distinct but coexisting kinds of meanings interplay within any text. The three kinds of metafunctions are related to three corresponding situational variables that operate in all communicative contexts. Any context of situation can be described in terms of three main variables that are important in influencing the semiotic choices that are made – FIELD, TENOR, and MODE. In addition to the context of situation (register) stratum, there is the context of culture (genre) as a higher stratum. Also, SFL approaches text analysis from a paradigmatic perspective, which privileges the systemic approach where structural operations are explained as realizing systemic choices. So when we analyse text, we show the functional organization of its structure; and we show what meaningful choices have been made, each one being seen in the context of what might have been meant but was not (Halliday, 2004). As a semiotic resource, language possesses an expression and a content plane and systems operating on each plane (Table 1-a). Thus, SFL allows texts to be viewed from above, below and around in understanding what meaningful choices have been made in the production of the text. It is this approach to text that allows SFL to be so readily adapted to the analysis of modalities other than verbal language.

Furthermore, although Halliday focused on language, he was very clear that this was only one semiotic system among many other modes of meaning in any culture, which might include '...*both art forms such as painting, sculpture, music, the dance, and so forth, and other modes of cultural behavior that are not classified under the heading of forms of art, such as modes of exchange, modes of dress, structures of the family, and so forth. These are all bearers of meaning in the culture. Indeed we can define a culture as a set of semiotic systems, as a set of systems of meaning, all of which interrelate'* (Halliday&Hasan, 1985, p. 4).

It is the metafunctional aspect of SFL and its link to the situational variables of social contexts that have provided a common theoretical basis for the development of similar 'grammatical' descriptions of the meaning-making resources of other semiotic modes. It is argued that these are aspects of meaning-making that apply to all semiotic resources (Lemke, 1998). These three types of meaning-making are inherent in all instances of communication, regardless of whether the communication is via language, image, music, sculpture or some other semiotic mode. Foregrounding the notion of meaning potential and choice in system networks, along with the principles of stratification and constituency, theorists working with modes other than language have effectively established a mapping of

the SFL metafunctions across modalities. With slight differences in nomenclature, the equivalent of Halliday's three metafunctional dimensions has been adapted to other social semiotics.

2.2. Systemic Functional-Multimodal Discourse Analysis

Multimodality is defined as the diverse ways in which multiple semiotic resources (language, visual images or sound, etc) are both co-deployed and co-contextualized in creating meaning. In its most basic sense, multimodality is the mixture of textual, audio, and visual modes in combination with mediums and materiality to produce meanings (Murray, 2013; Thibault, 2001; cited in Xiqin Liu and Dianning Qu, 2014). Meaning-making in mutimodal texts is dependent on an interplay of all semiotics resources of which language is only one component. As such, SF-MDA is 'an analytic practice which tests the application of the key principles of Systemic Functional Linguistics to the analysis of semiotic systems other than language and their interaction with each other and with language in semiosis' (Djonov, 2005, p. 73, cited in Lim Fei Victor, 2011).

The SF-MDA approach to multimodal texts involves the formulation of hierachies such as Items and Components where "larger-scalar units provide integrating contents for smaller-scale ones" (Baldry and Thibault, 2006, p. 144, cited in O'Halloran, 2008). The SF model for visual images based on O'Toole (1994) is displayed in Table 1-b. Parallel to language, it possesses a content plane (i.e. visual discourse/grammatical systems for the whole image and its component parts) and an expression plane (i.e. systems for the material realization of the image).

Given the constraints of an article, it is impossible to circumnavigate the numerous complexities in the different dimensions of multimodality. In the following sections, accounts of the three metafunctional potential and realisation systems will be confined to visual imagery in children's books, which serve as the framework to which the comparative description detailed in section 2 can be referred.

2.3. Reading visual imagery in children's books

Painter, Martin & Unsworth's (2013) recent study of children's picture books proposes networks of visual resources construing the three metafunctions, and a framework of how verbal and visual semiotic codes collaborate with each other to make meanings in narratives. This work is an extension from Kress and Van Leeuwen's (1996/2006) framework. According to the authors, "*While Kress and Van Leeuwen's (1996/2006) pioneering visual grammar provides an invaluable foundation for the understanding of a broad range of images, it is insufficiently developed for addressing key aspects of picture books – for example, the nature of relations between images in a sequence, the range of possibilities for point of view in a visual narrative and the visual resources for emotional engagement with the reader in such texts*." (p. 3). So, their work is "*designed particularly for the register of children's picture books, rather than images in general*" (ibid, p. 36).

The proposed visual systems of Painter, Martin & Unsworth's (2013) study offer a foundation for our systemic analysis of the illustrations in two sets of picture books for children in English and Vietnamese. The descriptive categories with systems of delicacy selections to construe the interpersonal, ideational and textual meanings are respectively summarized in Table 2, 3, and 4. Some elements which are not pertinent to our comparative analysis are deliberately left out for the sake of conciseness.

3. Illustrations in Picture books of Nursery Rhymes: A comparative analysis

The data for this study are the picture books of nursery rhymes in English and Vietnamese. The Vietnamese data are the ten volumes entitled *Đồng dao dành cho trẻ mầm non* (*Nursery Rhymes for Pre-school Children*; hereafter NRPC), published by Hong Bang Publisher, 2011. The English data are one hundred rhymes in *Mary Engelbreit's Mother Goose* (hereafter MG), published by Harpercollins, 2005. The decision to study these children's picture books is two fold: firstly, this genre is a significant site for a preliminary investigation into MDA, given its self-contained and concise nature; and secondly, due to its significant educational value, the findings may yield fruitful and pedagogically useful directions for further research in the same line.

Table 2. Meaning potential and realizations of interpersonal meaning

	Meaning potential	Realizations		
Affiliation	Focalization	contact (+ gaze)	direct (gazing participant faces viewer front on)	
			invited (gazing participant turns head/eyes to face viewer)	
		observe (- gaze)		
	Pathos	drawing style	appreciative (minimalist style)	
			empathic (generic style)	
			personalizing (naturalistic style)	
	Power	Vertical angle of viewing by viewer (high, mid, or low), by depicted participants in relation to another		
	Social distance/ proximity	Shot size; proximity/ touch of depicted participants		
	Involvement/ orientation	Horizontal angle of viewer; horizontal angle of character to other depiction; +/- mutuality of character gaze		
Feeling	Ambience	Color choices in relation to vibrancy, warmth and familiarity		
	Visual affect	Emotion depicted in facial features and bodily stance		
	Graduation: force	'Exaggerated' size, angle, proportion of frame filled, repetition of element		

Table 3. Meaning potential and realizations of ideational meaning

	Meaning potential	Realizations
Action	Visual action	Depicted action
	Action	Vectors
	Perception	Gaze vectors
	Cognition	Thought bubbles, face/hand gestures
	Talking	Speech bubbles, face/hand gestures
	Inter-event relations	Justaposition of images (+/– change of setting or participant)
Character	Character attribution	Depiction of physical attributes
	Character manifestation and appearance	Character depiction
	Character relations	Adjacent/symmetrical arrangement of different participants
Setting	Circumstantiation	Depiction of place, time, manner
	Inter-circumstance	Shifts, contrasts, continuities in locations

Table 4. Meaning potential and realizations of textual meaning

	Meaning potential	Realizations
Prominance	Framing	Binding of visual elements into units, separation of units via frames, margins, page edges
	Intermodal	Image and verbiage
	Integration	Placement within layout
	Focus	Compositional arrangement

As the two modals are intertwinedly linked in these books, we now present a very brief review concerning the verbal texts of these two sets of data before focusing on what and how the three meanings are conveyed in the images. Ideationally, Vo's (2016) study shows that Material and Relational processes are strongly deployed in MG, suggesting that the verses are largely concerned with actions and events that the characters get involved in, with the participants ranging from human beings to animals, from things to personified characters, from food and drinks to natural power. The Circumstances are moderately used. The same ideational picture is depicted in NRPC, as shown in Huynh's (2014) research findings. Furthermore, in order to probe into the interactional meaning, and given the lack of research in this area so far, we conducted an analysis of the attitudinal resources in 50

rhymes randomly chosen from MG. The quantitative analysis reveals a very restricted use of interpersonal expressions, with a meager number of 38 words/expressions in a total of 1574 words, accounting for just below 3% (2.41). However, a close analysis of the attendant imagery yields insightful findings regarding both the three meanings of the images and the text-image correlations. Following is a summary of some most noticeable features from the findings.
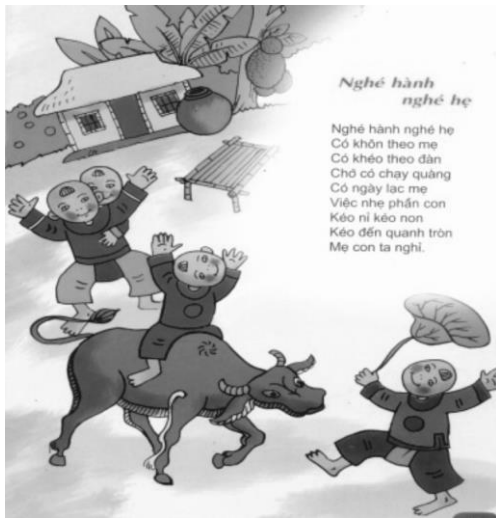

Figure 1.1. NR from NRPC, Vol. 4, p. 13

Figure 1.2. NR from NRPC, Vol. 3, p. 15

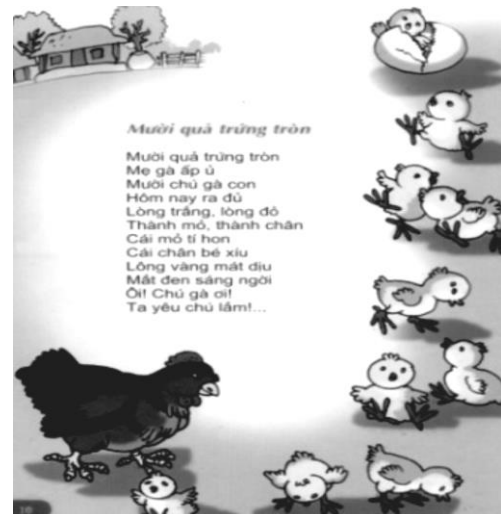Figure 1.3. NR from NRPC, Vol. 1, p. 11

Figure 1.4. NR from NRPC, Vo6. 1, p. 10

In terms of the three metafunctions, these two sets of data show the following resemblances. Compositionally, as can be seen from the sample figures 1.1-2.4, the verbal and visual components are made discrete sections. The 'complementary' layout is

dominant, suggesting that each has a distinct role to play in meaning making. Ideationally, both NRPC and MG faithfully depict the stories expressed in the verses, with the main participants made most salient. Circumstantial meaning in most images is prominently concerned with the spatial location for the depicted processes. Typically, the images specifying the details of the sky with clouds, the moon, the stars and/or the sun are common. With a common aim to familiarize the children with their own native culture, the children are presented with vast green meadows with sheep or horses in MG, and endless paddy fields with buffalos in NRPC. Another similarity lies with the abundant resources to represent interpersonal meanings. Close analysis indicates that it is the images that contribute to constructing evaluative stance as well as interaction; however, they are conveyed in different ways between the two cultures. There is also a consistent choice of the [warm] option of bright colors to make the images especially salient, aiming to entice and create a positive mood on the children readers.
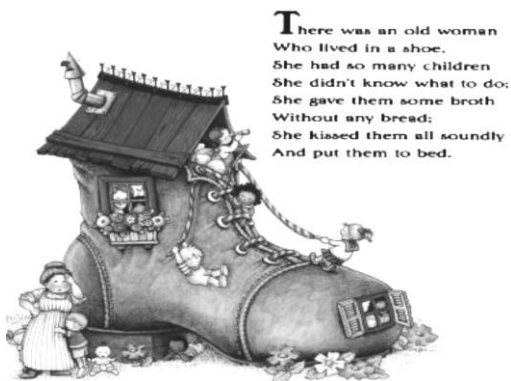
Figure 2.1. NR from MG, p. 29

Figure 2.2. NR from MG, p. 21

Figure 2.3. NR from MG, p. 66

Figure 2.4. NR from MG, p. 106

Apart from the similarities, close inspection indicates a number of differences. Compositionally, as far as the layout is concerned, the dominant ordering in MG is complementary: the pages are of a vertical placement with the illustrations being placed in

the lower or upper half of the page, as in Figures 2.1-2.4. There are also some cases for the verbiage and the image to be separated onto facing pages within a double-page layout, accounting for 15% of the data. Within the images, there are also frames created by the ideational content. However, this layout is not frequent in our data. By contrast, the choice of arrangement in NRPC is either integration, as in Figure 1.1 and 1.4 or complementary, such as Figure 1.2 and 1.4. Another noticeable difference is that a large number of the images in NRPC have a composition that divides them into sections with different ambience choices in the different parts. The salience of the text is shown by framing (as in Figure 1.3), which is a very frequent choice throughout the volumes. A split layout of this kind fails to make the image a coherent component as it is always the case in MG, as clearly shown in the figures. Furthermore, all the images in NRPC extend right to the page edges as unbound images; by contrast, in MG, there are always margins. The two sets also differ in the weight accorded to the images: whereas the two components tend to take up equal space in MG, images are privileged in relation to texts in NRPC, with the images in most cases taking up more space than the text blocks.

When it comes to ideational meaning, close inspection reveals delicate differences between the two cultures. As for NRPC, the minimalist style restricts variation in many features. For instance, head angles are with only front on and side views; delicate degrees of emotions are limited; the human participants feature ovals as heads and small black circles for eyes. Thus, on the pages we meet most participants whose gazes are directed at the viewer. For instance, in Figure 1.2, although the two characters face each other, the implied older sister's eye-gaze is not directed to her younger sister. We can also note that over the course of the volumes, some depicted participants are repeated irrespective of the different rhymes. In most cases, we are faced with the people in wooden slippers or bear-footed, dressed in patched traditional clothing, functioning as symbolic attributes suggesting the Vietnamese traditional culture. Also, in some cases, the 'peripheral' characters take up more attention, such as in Figure 1.1. Given the accompanying lads, the reader's attention to the main actor – the young buffalo – may be distracted.

On the contrary, MG features much more lifelikeness and detail; the consistent use of the generic style offers the potential for the greatest complexity and subtlety in the depiction of recognizable facial features and other elements. Contrasted with the white background, most images are salient with sharply drawn details and variation in vibrancy. Let's compare Figure 1.4 and Figure 2.2, which happen to be of the same ideational content. Figure 2.2 shows evidence that the artist focuses on every detail to depict the chickens in three distinct groups as conveyed in the text. Moreover, the reader's attention can easily be drawn to symbolic attributes of the hen in her caregiving role, gazing down at her preoccupied chickens lovingly from a high angle. Such meticulous attention can not be found in Figure 1.4, which is not only less vibrant with the [removed] option in using very few different colors but also less variations among the chicken.

The most salient difference between the two sets of data lies with the interactional meaning, conveyed via the expression of the emotions of the depicted participants and the interpersonal relations between them. As regard to *familiarity*, NRPC are more 'removed' than MG. The use of a reduced palette to signal a literal removal from reality can be seen in Figure 1.1 and Figure 1.4, where there are [removed] monochrome scenes in shades of all yellow and orange that contrast with the [familiar] more differentiated palette of the samples from MG. In Figure 2.1, for instance, the participants are depicted with enthusiasm and energy, which is realized by the use of a full range of colors – blue, green, yellow, pink, red, fawn, white and black. The exuberance is also emphasized by an upscaled graduation choice of mass – excessive size of the shoe to which many children are clinging.

In fact, most illustrations in MG create an emotional mood. With the generic style, the emotional repertoire is extended with more emotions enabled. For instance, in Figure 2.3, the girls are not only angled away from each other but also in a back-to-back orientation, suggesting the characters are disengaged and distant, something also indicated by their facial expressions. Besides, in a subdued, unsaturated background, the image suggests a feeling of sadness, while those depicting happiness and love in Figure 2.4 are in bolder, more highly saturated colors, encouraging us to share the joyousness of the moment. Furthermore, the images in MG consistently depict one participant as looking down or up the other participant from a higher or lower vertical angle, positioning them as having higher or lower power, as the woman with her children in Figure 2.1 or the hen with her chicken in Figure 2.2. By contrast, we are presented with participants of equal power in Figure 2.3 and 2.4. Gestures and bodily stance are also keys in the representation of affect. The depicted participants are sometimes shown in an intimate or close personal relationship, very often involving their bodies touching, such as Figure 2.1, 2.3, and 2.4. The participants and circumstances in a sense blend with each other, which conveys a sense of visual unity and lends compositional support to elaborate the ideational meaning of the texts as well.

4. Conclusion and implications

As Painter, Martin & Unsworth (2013) put it, "*while it is through the mediation of speaking adults that children are introduced to the verbal voice of a picture book, it is probably the visual images that are the most significant means for setting up an affective relationship between child and book, an important step in coming to terms with the print medium*" (p. 15). It is clear from our brief discussion that a huge contribution to meanings can be made by the illustrations in these children picture books. Especially, the detailed fine-grained analysis of the samples show that there are evidence that efforts have been made by both groups of artists to provoke an attitudinal response from the children readers, seemingly as an attempt to compensate for the relatively neutral ideational contents of the rhythms in both languages.

From the comparative findings, this study also addresses the urgent need for visual literacy and the lack of research into MDA in Vietnam. Findings of the study are in lines with the results of previous research of Unsworth and Ngo (2015) in terms of the concern over an effective manipulation of images as a meaning-making resource in the Vietnamese context. SF-MDA has proved to be a fruitful field of study. It is therefore necessary to conduct more research into the nature of the interactions between the verbal and the visual in various kinds of genres, particularly those of significant educational values, such as school science textbooks, as it has been focused in other countries.

References

Halliday, M. A. K., & Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective.* Geelong, VIC: Deakin University Press.

Halliday, M. A. K. (1994). *An introduction to functional grammar*. London: EdwardArnold.

Halliday, M. A. K. (2004). The language of science. In J. Webster (Ed.), *The collected works, 4.* London: Continuum.

Huynh, L. T. H. (2014). *An investigation into the lexical cohesion in English and Vietnamese nursery rhymes*. Unpublished M.A. thesis, Quy Nhon University.

Ink, M. E. (2005). *Mary Engelbreit's Mother Goose – One hundred best-loved verses*. New York: Harpercollins Publishers.

Kress, G., & van Leeuwen, T. (1996). *Reading images – The grammar of visual design*. New York: Routledge.

Lemke, J. ()). Multiplying meaning: Visual and verbal semiotics in scientific text, in J. R. Martin & R. Veel (Eds.), *Reading science – Critical and functional perspectives on discourses of science* (pp. 87-113). London: Routledge.

Lim, F. V. (2011). *A systemic functional multimodal discourse analysis approach to pedagogic discourse.* Unpublished doctoral thesis, National University of Singapore.

Lu. X., & Dianning Q. (2014). Exploring the multimodality of EFL textbooks for Chinese college students: A comparative study. *RELC Journal*, *45*(2), 135–150.

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

O'Halloran, K. L. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication, 7*(4), 443-475.

O'Toole, M. (1994). *The language of displaced art*. Leicester: Leicester University Press.

Painter, C., Martin, J., & Unsworth, L. (2013). *Reading visual narratives – Image analysis of children's picture books*. London: Equinox Publishing Ltd.

Quyen, D. (2011). *Đồng dao dành cho trẻ mầm non*. Hong Bang Publisher.

Ton. N. M. N. (2013). Diễn ngôn toán như một thể loại đa tín hiệu. *Ngôn ngữ*, số 4.

Ton. N. M. N. (2015). Phân tích diễn ngôn đa thức: Một số vấn đề lý luận và thực tiễn. Kỷ yếu *Hội*

thảo Khoa học Ngữ học Toàn quốc. Đại học Sài Gòn, tr. 227-236.

Unsworth, L. (2006). Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction. *English Teaching: Practice and Critique, 5*(1)*,* 55-76.

Unsworth, L., & Ngo, T. B. T. (2015). "Vai trò của hình ảnh trong sách giáo khoa dạy tiếng Anh ở Việt Nam. *T/c Ngôn ngữ & đời sống*, *1*(231), 93-99.

Vo, T. M. T. (2016). English nursery rhymes from a multimodal discourse analysis perspective. Unpublished M.A. Thesis, Quy Nhon University.

ĐỒNG DAO: NGHIÊN CỨU SO SÁNH TỪ CƠ SỞ PHÂN TÍCH DIỄN NGÔN ĐA THỨC - CHỨC NĂNG HỆ THỐNG

Tóm tắt. Trong giao tiếp ngày nay, người ta sử dụng không chỉ ngôn ngữ mà còn cả các phương tiện khác. Tuy nhiên, diễn ngôn đa thức vẫn còn là một lĩnh vực nghiên cứu tương đối mới mẻ. Mục đích của bài viết này là đóng góp vào hướng vào nghiên cứu mới này, và giới hạn trong phạm vi lý thuyết Phân tích diễn ngôn đa thức – chức năng hệ thống. Trước hết, bài viết tóm tắt các khía cạnh cơ bản của lý thuyết Chức năng hệ thống và của đường hướng Phân tích diễn ngôn đa thức – chức năng hệ thống. Sau đó, dựa trên khung phân tích hình ảnh trong sách hình dành cho thiếu nhi, bài viết trình bày một số kết quả so sánh đối chiếu thể loại đồng dao trong hai ngôn ngữ Anh và Việt. Cuối cùng là một số gợi ý ứng dụng với thực tiễn giáo dục ở Việt Nam.

Từ khóa: lý thuyết chức năng hệ thống, phân tích diễn ngôn, phân tích diễn ngôn đa thức