

## ĐỌC HIỂU ĐA PHƯƠNG THỨC TRUYỆN THIẾU NHI TRỰC TUYẾN: ĐỀ XUẤT KHUNG NĂNG LỰC DỰA TRÊN LÝ THUYẾT NGÔN NGỮ HỌC CHỨC NĂNG HỆ THỐNG VÀ NGỮ PHÁP HÌNH ẢNH

**TÔ KHÁNH DUY\***

**Abstract:** competency framework for multimodal reading comprehension grounded in Systemic Functional Linguistics (SFL) and Visual Grammar. Drawing on Halliday's metafunctional framework and Kress and van Leeuwen's visual grammar, the study investigates how language and images jointly construct ideational meaning in children's online narratives. Through an analysis of the transitivity system, representational structures, and language-image relations, the study identifies three core competencies required for learners: linguistic decoding, visual literacy, and multimodal integration. The findings contribute to the development of multimodal literacy education and provide pedagogical implications for curriculum design and learner support in contemporary educational contexts. The proposed competency framework may assist teachers in designing instructional activities that enable learners to interpret, connect, and evaluate information across multiple semiotic modes.

**Keywords:** *multimodal discourse, online children's stories, Systemic Functional Linguistics, visual grammar, multimodal reading comprehension.*

### 1. Giới thiệu

Trong bối cảnh giáo dục đương đại, năng lực đọc hiểu không còn giới hạn ở việc giải mã văn bản thuần ngôn ngữ mà được mở rộng thành khả năng tương tác và diễn giải các văn bản đa phương thức. Người học ngày nay tiếp xúc thường xuyên với các sản phẩm số như truyện tranh trực tuyến, sách nói có hình ảnh, video kể chuyện,..., nơi ngôn ngữ và hình ảnh kết hợp để kiến tạo ý nghĩa. Sự thay đổi này đòi hỏi giáo dục phải chuyển dịch từ “năng lực đọc hiểu (literacy)” truyền thống sang “năng lực đọc hiểu đa phương thức (multimodal literacy)” - năng lực tiếp nhận, phân tích và tích hợp thông tin từ nhiều phương thức kí hiệu khác nhau.

Truyện thiếu nhi trực tuyến là một dạng văn bản đa phương thức điển hình, kết hợp giữa văn bản ngôn ngữ (text), hình ảnh (image), và trong một số trường hợp là âm thanh (soundscape) hoặc yếu tố tương tác (interactive elements). Tuy nhiên, hiện nay, việc phát triển khung năng lực đọc hiểu đa phương thức truyện thiếu nhi trực tuyến tiếng Việt còn hạn chế.

Dựa trên lý thuyết Ngôn ngữ học Chức năng Hệ thống (Halliday 1978, 1985, 1994; Halliday & Matthiessen 2004, 2014) và Ngữ pháp hình ảnh (Kress & van Leeuwen, 2006), bài báo này đề xuất một khung năng lực đọc hiểu đa phương thức dành cho thiếu nhi, đặt nền tảng trên mối quan hệ giữa ngôn ngữ - hình ảnh trong truyện thiếu nhi trực tuyến.

### 2. Ngôn ngữ học chức năng hệ thống và Ngữ pháp hình ảnh

Ngôn ngữ học Chức năng Hệ thống là lý thuyết ngôn ngữ học được phát triển bởi Michael Halliday (Halliday, 1978, 1985, 1994; Halliday & Matthiessen, 2004, 2014). Lý thuyết này được cho là chịu ảnh hưởng bởi các công trình của J.R.Firth (1890 - 1960), Malinowski (1844 - 1942). Ngôn ngữ học Chức năng Hệ thống chú trọng vai trò mạnh mẽ của ngôn ngữ trong cuộc sống và xem việc tạo ra nghĩa là một quá trình mà qua đó ngôn ngữ định hình và được định hình bởi các bối cảnh mà nó được sử dụng. Mỗi ngôn ngữ đều cung cấp cho người nói/người viết vô số lựa chọn để giải thích nghĩa. Vì thế, theo Ngôn ngữ học Chức năng Hệ thống, ngôn ngữ được xem là *hệ thống những sự lựa chọn (system of choices)*.

\* Trường Đại học Khoa học Xã hội và Nhân văn, ĐHQG Thành phố Hồ Chí Minh; Email: tokhanhduy@hcmussh.edu.vn

Ngôn ngữ học Chức năng Hệ thống đề cập đến ba chức năng/ngĩa trừu tượng (còn gọi là *siêu chức năng* - metafunction), chúng được thực hiện đồng thời trong mọi cú (clause) mà chúng ta nói hoặc viết, và liên hệ các lựa chọn ngôn ngữ của chúng ta với chu cảnh mà ngôn ngữ tham gia. Ba siêu chức năng bao gồm: *Ý niệm* (ideational), *Liên nhân* (interpersonal) và *Văn bản* (textual). Ba siêu chức năng này được thể hiện như sau:

*Siêu chức năng Ý niệm (ideational metafunction)*: thể hiện kinh nghiệm của con người về thế giới vật chất (bao gồm các sự vật, sự kiện, và mối quan hệ giữa chúng) và cả thế giới nội tâm. Cách diễn đạt này sử dụng các thành phần ngữ pháp như quá trình (process), tham thể (participant), và chu cảnh (circumstance).

*Siêu chức năng Liên nhân (interpersonal metafunction)*: xây dựng và duy trì mối quan hệ giữa người nói/người viết và người nghe/người đọc. Chức năng này thể hiện qua thức (mood), tình thái (modality) và vai trò của tham thể trong giao tiếp.

*Siêu chức năng Văn bản (Textual metafunction)*: tạo sự mạch lạc và tổ chức cho văn bản, đảm bảo rằng ngôn ngữ được sử dụng một cách trôi chảy và phù hợp với bối cảnh. Các yếu tố như Đề (theme) và Thuyết (rheme) đóng vai trò quan trọng trong tổ chức thông tin. Siêu chức năng văn bản chính là cách thức thể hiện, sự tổ chức, vật chất hóa hai siêu chức năng còn lại của ngôn ngữ để hình thành một thông điệp hoàn chỉnh.

*Ngữ pháp hình ảnh* có tên tiếng Anh là “Reading Images: The Grammar of Visual Design” được xem như một trong những “phiên bản” của Ngôn ngữ học Chức năng Hệ thống, như một công cụ để phân tích các mô hình hình kí hiệu có nguồn tạo kí hiệu là phương thức hình ảnh (ngoài phương thức ngôn ngữ). Công trình này đặt mục tiêu phát triển một “ngữ pháp hình ảnh” (*visual grammar*), phân tích cách các yếu tố hình ảnh như màu sắc, bố cục, phối cảnh, và các mối quan hệ không gian phối hợp để tạo ra nghĩa.

Ba loại nghĩa trong Ngữ pháp hình ảnh được phát triển từ ba loại nghĩa trong Ngôn ngữ học Chức năng Hệ thống.

| Ba loại nghĩa trong Ngôn ngữ học chức năng hệ thống |                                                                                     | Ba loại nghĩa trong Ngữ pháp hình ảnh       |
|-----------------------------------------------------|-------------------------------------------------------------------------------------|---------------------------------------------|
| Nghĩa Ý niệm<br>(Ideational meaning)                |  | Nghĩa Biểu hiện<br>(Representation meaning) |
| Nghĩa Liên nhân<br>(Interpersonal meaning)          |                                                                                     | Nghĩa Tương tác<br>(Interaction meaning)    |
| Nghĩa Văn bản<br>(Textual meaning)                  |                                                                                     | Nghĩa Tổ chức<br>(Composition meaning)      |

### 3. Phương thức và diễn ngôn đa phương thức

Phương thức là một hệ thống kí hiệu được xã hội thừa nhận và sử dụng ổn định để tạo nghĩa trong giao tiếp. Theo Kress (2014) một nguồn tạo kí hiệu (semiotic resource) chỉ được xem là *phương thức* khi nó được cộng đồng sử dụng thường xuyên, có quy ước xã hội, và đáp ứng được 3 siêu chức năng của Halliday: Ý niệm (biểu hiện nội dung, sự kiện, kinh nghiệm); Liên nhân (thái độ, quan hệ xã hội) và Văn bản (tính mạch lạc, tổ chức thông điệp).

Theo Arola, Sheppard & Ball (2014) và Jewitt (2014a), có 5 phương thức phổ biến: (1) Phương thức ngôn ngữ (Linguistic mode) Bao gồm: từ ngữ, câu, đoạn, văn bản nói hoặc viết; thể hiện ý tưởng, thông tin, mạch lạc, cấu trúc câu chuyện. (2) Phương thức hình ảnh (Visual mode), bao gồm: màu sắc, bố cục, đường nét, phong cách, góc nhìn, kích thước; truyền đạt nhanh và trực tiếp bằng thị giác; đặc biệt quan trọng trong truyện tranh. (3) Phương thức âm thanh (Aural mode), Gồm: nhạc nền, tiếng động, giọng nói, độ cao - thấp, âm sắc. Hiện diện trong video, podcast, clip minh họa. (4) Phương thức không gian (Spatial mode), Liên quan đến bố trí, tổ chức, khoảng cách giữa các đối tượng trên trang/khung hình; Quyết định tính nhấn mạnh, trọng tâm, thứ bậc thông tin. (5) Phương thức cử chỉ

(Gestural mode), Gồm: biểu cảm khuôn mặt, cử động tay, tư thế, tương tác cơ thể. Quan trọng trong giao tiếp trực tiếp, sân khấu, video hoạt hình.

Đặc điểm của các phương thức là bình đẳng trong việc tạo nghĩa: không chỉ ngôn ngữ mà mọi phương thức đều có vai trò. Chúng không hoạt động đơn lẻ: nghĩa của diễn ngôn đa phương thức được tạo ra từ sự phối hợp giữa các phương thức. Các phương thức cũng bị chi phối bởi văn hóa - xã hội: cách dùng mỗi phương thức phụ thuộc vào cộng đồng và bối cảnh sử dụng.

Khái niệm phương thức dần trở nên phổ biến khi nghiên cứu về diễn ngôn đa phương thức trở thành một hướng nghiên cứu đầy tiềm năng. *Diễn ngôn đa phương thức* (multimodal discourse) là diễn ngôn có sự góp mặt nhiều hơn một phương thức để tạo nghĩa. Theo Arola, Sheppard & Cheryl E. Ball (2014), từ “multimodal” là sự kết hợp của “multiple” và “mode”. Jewitt (2014b) lập luận rằng đa phương thức mở rộng trọng tâm của sự tương tác, khám phá cách mọi người sử dụng cử chỉ, ánh mắt, tư thế, chuyển động, không gian,... để làm trung gian cho sự tương tác trong một bối cảnh nhất định (Jewitt, 2014b: 37).

Trong thực tế sử dụng, ngôn ngữ và hình ảnh không được sử dụng độc lập mà thường là những yếu tố đi kèm nhau, giữa chúng tồn tại một mối quan hệ về logic. Hoàng Văn Vân (2025) cho rằng khi các cú đơn giản được liên kết hợp lí với nhau, thì việc phân tích văn bản cũng nên liên quan đến việc kiểm tra các cú kết hợp với nhau để tiết lộ cách văn bản trở thành một tổng thể thống nhất. Có thể thấy, các phương thức có sự khác biệt nhất định, vai trò của chúng trong diễn ngôn vì thế cũng không giống nhau.

Trong một nỗ lực chỉ ra sự kết hợp của ngôn ngữ và hình ảnh, O’Halloran (2008) cũng đã đưa ra được những đặc điểm tương quan giữa ngôn ngữ - hình ảnh dựa trên cơ sở phân tầng của Ngôn ngữ học Chức năng Hệ thống:

*Bảng 1. Mô hình tiếp cận ngôn ngữ và hình ảnh theo quan điểm của O’Halloran (2008)*

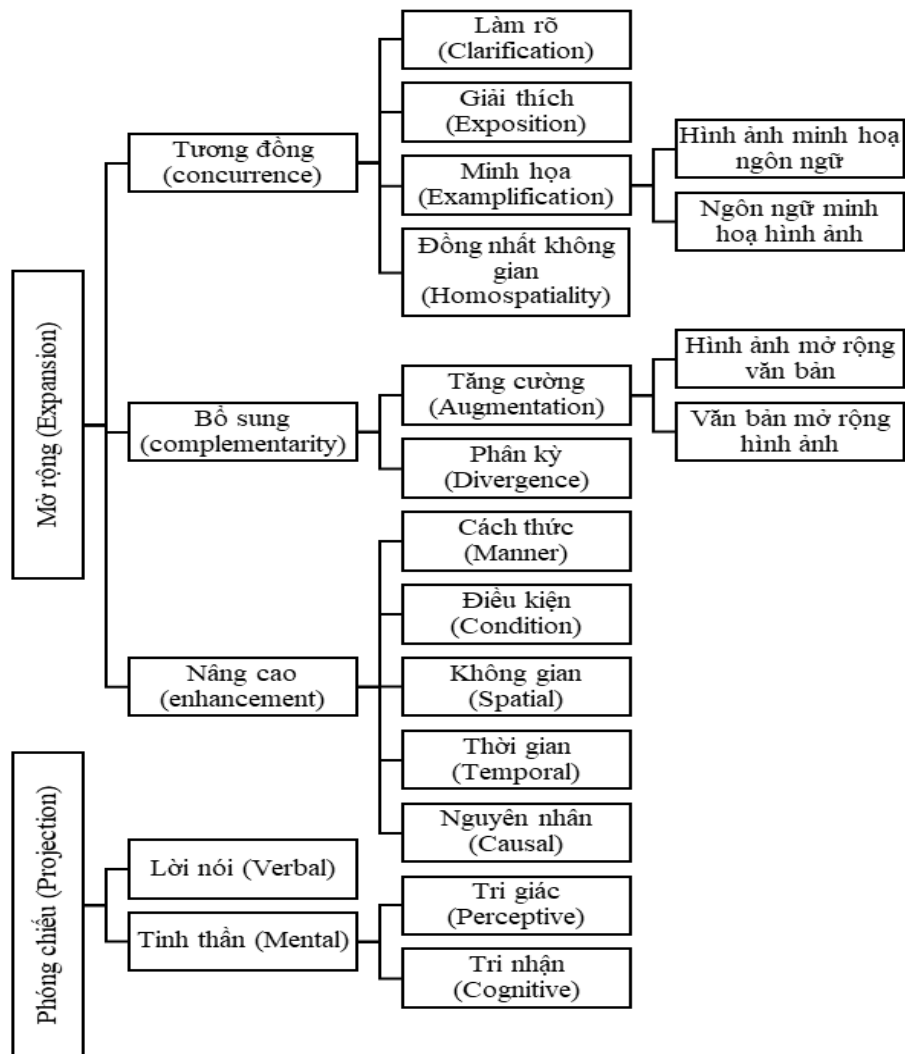
|                                        | NGÔN NGỮ                                                                                                                                                                                                                                      | HÌNH ẢNH                                                                                                                                                                                                                                                             |
|----------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Tầng nội dung<br>(Content stratum)     | Ngữ nghĩa diễn ngôn (Discourse semantics)<br>Các mối quan hệ trong diễn ngôn (đoạn, bài)<br>Từ ngữ - ngữ pháp<br>Phức hợp cú (Clause complex)<br>Cú (Clause)<br>Cụm từ (Word Group)<br>Từ (Word)<br>(Các hệ thống dựa vào các siêu chức năng) | Ngữ nghĩa diễn ngôn (Discourse semantics)<br>Những quan hệ giữa các bộ phận hình ảnh (Intervisual Relations)<br>Công trình (Works)<br>Ngữ pháp (Grammar)<br>Cảnh (Scence)<br>Đoạn (Episode)<br>Hình (Figure)<br>Phần (Part)<br>(Các hệ thống dựa vào siêu chức năng) |
| Tầng biểu hiện<br>(Expression stratum) | Kiểu in/ chữ viết và Ngữ âm (Typography/ Graphology and Phonology)<br>Các hệ thống giao chức năng (Cross-Functional Systems)                                                                                                                  | Hình (Graphics)<br>Các hệ thống giao chức năng (Cross-Functional Systems)                                                                                                                                                                                            |

**4. Đọc hiểu truyện thiếu nhi trực tuyến**

Đối tượng của bài báo này là sự phối hợp các phương thức trong việc xây dựng nghĩa ý niệm trong truyện thiếu nhi trực tuyến. Dựa trên đối tượng này, chúng tôi nghiên cứu trên 30 truyện thiếu nhi trên nền tảng truyện thiếu nhi trực tuyến Thư viện Mây (Literacy Cloud). Kết quả nghiên cứu sẽ chỉ ra được mối quan hệ giữa các phương thức trong truyện thiếu nhi trực tuyến đồng thời đây sẽ là cơ sở để bài báo đề xuất khung năng lực đọc hiểu đa phương thức.

Mô hình mối quan hệ giữa ngôn ngữ - hình ảnh do Len Unsworth (2006) đề xuất, giúp chỉ ra các mối quan hệ tiềm năng của ngôn ngữ và hình ảnh trong việc xây dựng nghĩa ý niệm của diễn ngôn đa phương thức.

Các phương thức không chỉ được đóng khung trong cùng một miền mà chúng hòa vào nhau cùng nhau xây dựng diễn ngôn. Do đó, việc đọc hiểu diễn ngôn đa phương thức không phải là đọc hiểu từng diễn ngôn riêng lẻ mà là đọc hiểu sự phối hợp các phương thức trong diễn ngôn.



Sơ đồ Mô hình mối quan hệ ngôn ngữ - hình ảnh của Len Unsworth (2006)

Bảng 2. Kết quả khảo sát mối quan hệ ngôn ngữ - hình ảnh trên 30 truyện thiếu nhi trực tuyến thuộc nền tảng truyện trực tuyến Thư viện Mây

| Kiểu quan hệ | Loại quan hệ | Số lượng (lượt) | Tỉ lệ (%) |
|--------------|--------------|-----------------|-----------|
| Mở rộng      | Tương đồng   | 347             | 47,4      |
|              | Bổ sung      | 122             | 16,7      |
|              | Nâng cao     | 64              | 8,7       |
| Đôi chiếu    | Bằng lời     | 173             | 23,6      |
|              | Tinh thần    | 26              | 3,6       |
| Tổng         |              | 732             | 100,0     |

Kết quả nghiên cứu thể hiện rõ tiềm năng kết hợp đa dạng của ngôn ngữ và hình ảnh trong diễn ngôn truyện thiếu nhi trực tuyến. Trong đó, kiểu *mở rộng* (72,8%) phổ biến hơn kiểu *phóng chiếu* (27,2%). Trong *mở rộng*, kiểu *tương đồng* phổ biến nhất với 47,4%, tiếp theo là kiểu *bổ sung* với 16,7 % và ít nhất là kiểu *nâng cao* với 8,7%. Trong kiểu *phóng chiếu*, kiểu *phóng chiếu bằng lời* chiếm ưu thế với 23,6% so với 3,6 % của kiểu *phóng chiếu tinh thần*.

Kết quả nghiên cứu cung cấp một cách nhìn về mối quan hệ ngôn ngữ và hình ảnh trong diễn ngôn truyện thiếu nhi. Chúng tôi nhận thấy rằng ngôn ngữ và hình ảnh không tồn tại một cách độc lập

trong diễn ngôn mà chúng phối hợp với nhau để chuyển tải nghĩa cho diễn ngôn. Bên cạnh đó, sự phối hợp giữa ngôn ngữ và hình ảnh là tương tác hai chiều trong việc kiến tạo nghĩa cho diễn ngôn.

### 5. Khung năng lực đọc hiểu đa phương thức

Dựa trên đặc điểm cấu trúc của truyện thiếu nhi trực tuyến, cũng như kết quả phân tích ngôn ngữ, hình ảnh và quan hệ ngôn ngữ - hình ảnh, nghiên cứu đề xuất một khung năng lực đọc hiểu đa phương thức gồm ba nhóm năng lực cốt lõi: giải mã ngôn ngữ, hiểu biết thị giác và tích hợp đa phương thức. Ba nhóm năng lực này được nhìn nhận như những thành tố bổ trợ lẫn nhau, tạo nên nền tảng cho việc tiếp nhận và diễn giải ý nghĩa trong các văn bản đa phương thức dành cho học sinh tiểu học.

#### a. Năng lực giải mã ngôn ngữ

Trước hết, năng lực giải mã ngôn ngữ (linguistic decoding) là năng lực nền tảng, cho phép học sinh nhận diện và hiểu các thành tố tạo nghĩa trong câu, bao gồm loại quá trình, người tham gia và hoàn cảnh theo phân loại của Halliday. Khi học sinh nhận biết được ý nghĩa của hành động, trạng thái hoặc cảm xúc qua cấu trúc chuyên tác, học sinh có thể hình dung rõ ràng hơn về diễn biến sự kiện và mối quan hệ giữa các nhân vật trong truyện. Khả năng này không chỉ giúp học sinh theo dõi mạch truyện mà còn là tiền đề để người học đối chiếu và liên kết thông tin ngôn ngữ với các phương thức khác.

#### b. Năng lực hiểu biết thị giác

Tiếp theo, năng lực hiểu biết thị giác (visual literacy) đóng vai trò trung tâm trong tiếp nhận truyện thiếu nhi trực tuyến, bởi hình ảnh là phương thức chiếm ưu thế về mật độ thông tin và sức gợi cảm. Năng lực này bao gồm khả năng nhận diện nhân vật, bối cảnh, hướng vector hành động, biểu cảm khuôn mặt, bố cục khung hình cũng như các yếu tố thị giác như màu sắc, kích thước hay vị trí. Khi học sinh hiểu được chức năng tạo nghĩa của các yếu tố thị giác, học sinh có thể diễn giải hành động, cảm xúc hoặc bầu không khí của câu chuyện ngay cả khi văn bản ngôn ngữ không miêu tả trực tiếp. Điều này giúp hình ảnh phát huy đúng vai trò như một nguồn tạo kí hiệu độc lập trong quá trình kiến tạo nghĩa cho diễn ngôn.

#### c. Năng lực tích hợp đa phương thức

Cuối cùng, năng lực tích hợp đa phương thức (multimodal integration) là năng lực mang tính chất tổng hợp cao nhất, thể hiện khả năng liên kết, so sánh và đánh giá thông tin đến từ văn bản ngôn ngữ và hình ảnh. Học sinh cần nhận biết được sự tương ứng giữa hai phương thức, đồng thời xác định phần thông tin được mở rộng, bổ sung hoặc nâng cao ý nghĩa nhờ yếu tố hình ảnh. Năng lực tích hợp còn bao gồm khả năng phát hiện sự thiếu thống nhất giữa nội dung văn bản và minh họa, cũng như khả năng rút ra thông tin suy luận từ sự kết hợp đó. Đây chính là năng lực then chốt giúp người học đọc hiểu truyện trực tuyến một cách sâu sắc và toàn diện, thay vì chỉ tiếp nhận thông tin theo từng kênh riêng lẻ.

Nhìn chung, ba năng lực trên không tồn tại tách biệt mà tương tác mật thiết trong quá trình đọc hiểu truyện thiếu nhi trực tuyến. Việc trang bị cho học sinh khung năng lực đọc hiểu đa phương thức này giúp các em phát triển khả năng diễn giải, kết nối và đánh giá thông tin một cách linh hoạt trong môi trường số hiện đại, nơi văn bản đa phương thức ngày càng chiếm ưu thế. Khung năng lực cũng tạo cơ sở để giáo viên thiết kế các hoạt động dạy học phù hợp, góp phần nâng cao năng lực tiếp nhận và tư duy phản biện của học sinh tiểu học.

### 6. Kết luận và đề xuất

Nghiên cứu này đã tiếp cận truyện thiếu nhi trực tuyến như một hình thức diễn ngôn đa phương thức điển hình, trong đó ngôn ngữ và hình ảnh cùng tham gia kiến tạo nghĩa ý niệm. Trên cơ sở lí thuyết Ngôn ngữ học Chức năng Hệ thống và Ngữ pháp hình ảnh, bài báo đã phân tích đặc điểm của từng phương thức cũng như mối quan hệ giữa các phương thức trong quá trình tạo nghĩa, qua đó làm rõ cơ chế vận hành của diễn ngôn truyện thiếu nhi trong môi trường số.

Kết quả phân tích cho thấy ngôn ngữ và hình ảnh không tồn tại như những kênh truyền tải thông tin tách biệt mà liên kết chặt chẽ với nhau theo các kiểu quan hệ đa dạng, trong đó quan hệ mở rộng

chiếm ưu thế. Điều này cho thấy hình ảnh trong truyện thiếu nhi trực tuyến không chỉ có chức năng minh họa mà còn đảm nhiệm vai trò bổ sung, cụ thể hóa và nâng cao nội dung ý niệm được thể hiện qua ngôn ngữ. Sự phối hợp hai chiều giữa các phương thức góp phần tạo nên tính mạch lạc và chiều sâu ý nghĩa cho diễn ngôn, đồng thời phù hợp với đặc điểm tiếp nhận của người học ở lứa tuổi thiếu nhi.

Từ những phát hiện trên, nghiên cứu đã đề xuất một khung năng lực đọc hiểu đa phương thức gồm ba thành tố cốt lõi: năng lực giải mã ngôn ngữ, năng lực hiểu biết thị giác và năng lực tích hợp đa phương thức. Khung năng lực này được xây dựng dựa trên các nguồn lực tạo nghĩa chủ yếu của truyện thiếu nhi trực tuyến và phản ánh yêu cầu thực tế của hoạt động đọc hiểu trong môi trường học tập số. Ba nhóm năng lực không tồn tại độc lập mà có mối quan hệ hỗ trợ lẫn nhau, giúp người học tiếp nhận, diễn giải và đánh giá thông tin một cách toàn diện hơn so với cách tiếp cận đọc hiểu truyền thống.

Về mặt lí luận, nghiên cứu góp phần mở rộng hướng tiếp cận Ngôn ngữ học Chức năng Hệ thống sang phân tích diễn ngôn đa phương thức trong tiếng Việt, đồng thời khẳng định tính khả dụng của Ngữ pháp hình ảnh trong nghiên cứu truyện thiếu nhi trực tuyến. Về mặt thực tiễn, khung năng lực đọc hiểu đa phương thức được đề xuất có thể được xem là cơ sở tham chiếu cho việc thiết kế hoạt động dạy học, xây dựng tài liệu và đánh giá năng lực đọc hiểu của học sinh tiểu học trong bối cảnh giáo dục hiện đại.

Tuy nhiên, do phạm vi nghiên cứu còn giới hạn ở một số lượng truyện nhất định và mới tập trung chủ yếu vào hai phương thức ngôn ngữ và hình ảnh, các nghiên cứu tiếp theo có thể mở rộng dữ liệu, đồng thời xem xét vai trò của các phương thức khác như âm thanh, không gian và yếu tố tương tác. Việc khảo sát trực tiếp quá trình tiếp nhận của học sinh cũng là một hướng nghiên cứu cần thiết nhằm kiểm chứng và hoàn thiện hơn khung năng lực đọc hiểu đa phương thức đã được đề xuất.

#### NGUỒN NGŨ LIỆU

1. Thư viện Mây (Literacy Cloud). <https://literacycloud.org/>

#### TÀI LIỆU THAM KHẢO

1. Halliday, M. A. K. (1978). *Language as Social Semiotic: The Interpretation of Language and Meaning*. London: Edward Arnold.
2. Halliday, M. A. K. (1985). *An Introduction to Functional Grammar (First edition)*. London: Edward Arnold.
3. Halliday, M. A. K. (1994). *An Introduction to Functional Grammar (Second edition)*. London: Edward Arnold.
4. Halliday, M.A.K & Matthiessen, M.I.M. (2004). *Halliday's Introducing to Functional Grammar (Third edition)*. Edward Arnold.
5. Halliday, M.A.K & Matthiessen, M.I.M. (2014). *Halliday's Introducing to Functional Grammar (Fourth edition)*. London, New York: Routledge.
6. Kress, G & van Leeuwen, T. (2006). *Reading Image: The Grammar of Visual Design (Second edition)*. Routledge.
7. Van, Hoang. (2025). *Clause complex resources realising logical meanings in a literary text: a systemic functional linguistics analysis*. VNU Journal of Foreign Studies, Vol.41, 1-17.
8. O' Halloran, K. L. (2008). *Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery*. Visual Communication, Vol 7(4), 443-475. <https://doi.org/10.1177/1470357208096210>.
9. Kress, G. (2014). What is mode?. In Jewitt, C (Ed.) *The Routledge Handbook of Multimodal Analysis (Second edition)*, 60-75. London, New York: Routledge.
10. Arola, K. L. Sheppard, J., & Ball, C. E. (2014). *Writer/Designer: A guide to making multimodal projects*. New York: Bedford/St.Martin's.
11. Jewitt, C. (2014a). *An introduction to Multimodality*. In Jewitt, C (Ed.) *The Routledge Handbook of Multimodal Analysis (Second edition)*, 15-30). London, New York: Routledge.
12. Jewitt, C. (2014b). *Different approaches to multimodality*. In Jewitt, C (Ed.) *The Routledge Handbook of Multimodal Analysis (Second edition)*, 31-43. London, New York: Routledge.
13. Unsworth, L. (2006). *Image/Text Relations and Intersemiosis: Towards Multimodal Text Description for Multiliteracies Education*. Proceedings of the 33rd International Systemic Functional Congress, 1165-1205.