Sign Languages and Linguistics Convergence Research
Vol. 2. No. 2 (2025) pp.137~162

Visual Grammar as a Theoretical Framework for Sign Language Analysis: A Multimodal Study of Korean Sign Language

Lee, Seonhye · Dai, Ziyun · Jhang, Se-Eun*
(Pukyong National University · Korea Maritime & Ocean University · Korea Maritime & Ocean University)

Lee, Seonhye, Dai, Ziyun, & Jhang, Se-Eun. 2025. Visual Grammar as a Theoretical Framework for Sign Language Analysis: A Multimodal Study of Korean Sign Language. ***Sign Languages and Linguistics Convergence Research,*** *2*(2), 137-162. Sign languages are multimodal, integrating manual articulators such as handshape, movement, and location with non-manual elements including facial expression, gaze, and body posture to construct meaning across parallel channels. Although Kress and van Leeuwen's Visual Grammar (VG) has been extensively applied to static visual texts, its application to dynamic, time-based signed discourse remains underexplored. This study proposes an exploratory socio-semiotic framework that maps VG's three metafunctions—representational, interactive, and compositional meanings—onto coordinated manual, non-manual, and spatial–temporal resources in Korean Sign Language (KSL). The analysis draws on selected discourse excerpts from the National Institute of Korean Language KSL Nuri Dictionary, transcribed and time-aligned using ELAN, focusing on six focal lexical items: CAR, OIL, ENGINE, FOR LONG, CHANGE, and DANGER. Qualitative findings reveal recurring tendencies: manual parameters foreground representational meaning, non-manual features foreground interactive meaning, and spatial–temporal organization contributes to compositional meaning, while allowing functional overlap across semiotic channels. By operationalizing VG for embodied and temporally unfolding discourse, this study provides an analytic basis for extending the framework to sign languages

* First author: Lee, Seonhye, Corresponding author: Jhang, Se-Eun, Co-author: Dai, Ziyun

and motivates future cross-linguistic research using larger datasets.

Key Words: Visual Grammar, Korean Sign Language, Sign Language Discourse Analysis, Multimodality, Social Semiotics, ELAN

1. Introduction

Sign languages are inherently multimodal, integrating manual articulators (handshape, movement, and location) with non-manual elements (NMEs) such as facial expressions, gaze, and body posture to construct discourse. The simultaneous nature of these resources allows sign languages to express complex meanings across parallel channels. While this inherent multimodality has been acknowledged within sign linguistics, a systematic framework capable of analyzing the comprehensive semiotic structure of signed discourse remains a central theoretical challenge within sign linguistics.

The Visual Grammar (VG) framework, originally developed by Kress and van Leeuwen (2020) based on Halliday's Systemic Functional Linguistics (SFL), has proven highly effective for analyzing the social semiotics of static visual texts like images and advertisements. However, its systematic application to dynamic, time-based, and embodied communication systems—such as sign languages—has been largely unexplored. This gap stems from the conceptual difficulty of mapping VG's visual design principles onto the unique temporal and spatial organization of signed discourse.

This study addresses this critical research gap by proposing and validating a novel socio-semiotic framework that systematically maps the three metafunctions of VG—Representational, Interactive, and Compositional Meanings—onto the distinct multimodal constituents of Korean Sign Language (KSL). This study aims to demonstrate the analytic feasibility and explanatory potential of extending VG beyond static media to a fully structured visual language system.

To achieve this goal, this paper proposes an analytical framework and seeks to systematically answer the following research questions (RQs) through a qualitative case study of KSL, as shown in (1) below:

(1) a. **RQ1:** How is VG's Representational Meaning realized through the manual elements (handshape, movement, location) of KSL discourse?
b. **RQ2:** How is VG's Interactive Meaning realized through the NMEs (gaze, facial expression, body posture) of KSL discourse?
c. **RQ3:** How is VG's Compositional Meaning realized through the spatial and temporal organization of KSL discourse?

This study is designed as an exploratory case study based on a limited KSL dataset. Rather than advancing universal grammatical claims, it aims to examine the analytical applicability of VG to sign language discourse and to identify recurring patterns observable within the current data.

The remainder of this paper is structured as follows: Section 2 provides the theoretical background of SFL and VG. Section 3 details the methodology, including the data corpus and the conceptual framework used for analysis. Section 4 presents the findings, and Section 5 summarizes the conclusion, theoretical contributions, and implications for future research.

2. Theoretical Background

2.1. Systemic Functional Linguistics and Its Recontextualization in Visual Grammar

SFL posits that semiotic systems simultaneously realize three metafunctions: ideational (construing experience), interpersonal (enacting social relations), and textual (organizing coherent messages) (Halliday, 1985). Building on this framework, Kress and van Leeuwen (2006) extend the metafunctional principle to visual communication—not by mechanical transfer—but by reconceptualizing these functions in relation to the affordances of visual modes. Accordingly, within VG, the ideational, interpersonal, and textual metafunctions are rearticulated as representational, interactive, and compositional meanings, respectively, reflecting the distinct semiotic resources through which visual meaning is organized.

This theoretical framework is particularly well suited for the analysis of KSL, an inherently visual-spatial and embodied linguistic system. Previous sign language research has demonstrated that KSL constructs meaning through the coordinated deployment of manual articulators, non-manual markers, and the structured use of signing space (Johnston, 2013). From this perspective, the application of VG to KSL can be understood as a theoretically motivated extension of Hallidayan social semiotics to a modality in which visual organization is grammatically central.

By conceptualizing KSL as a multimodal semiotic system, the present study examines how representational, interactive, and compositional resources interact dynamically to form multidimensional meaning structures. This theoretical orientation establishes the conceptual foundation for the exploratory analyses presented in Sections 3 and 4 and is consistent with broader trends in multimodal discourse research, which conceptualize meaning as emerging from the coordinated deployment of multiple semiotic modes rather than from any single representational system (Anderson et al., 2022).

2.2. The Grammar of Visual Design

Kress and van Leeuwen (1996; 2006; 2020) translated Halliday's (1985) three metafunctions into a grammar for visual communication: The grammar of visual design. This framework articulates how images and visual compositions encode meaning by organizing semiotic resources into systematic categories. In VG, the metafunctional organization is realized through three major meaning systems: Representational Meaning, Interactive Meaning, and Compositional Meaning. The following subsections outline each of these metafunctions in detail.

2.2.1. Representational Meaning

This function relates to how the sign represents the world, constructing events, actions, and concepts. It is categorized into two processes as can be seen in (2) below.

(2) a. Narrative processes: Depicting actions or changes, involving participants

(actors, goals) and vectors (lines of force or movement). These processes are further classified as actional (unidirectional or bidirectional) or reactional (unidirectional or bidirectional) (Kress & van Leeuwen, 2006).

b. Conceptual processes: Depicting generalized categories, without vector lines. These include classificational (taxonomies), analytical (part-whole structures), or symbolic (representing identity, meaning, or value).

2.2.2. Interactive Meaning

This function defines the relationship between the visual producer (signer) and the viewer (addressee). Key parameters include the following three sub-domains as illustrated in (3).

(3) a. Gaze (demand vs. offer): Whether the signer's eyes directly address the viewer (demand) or look away (offer).

b. Distance: The implied social distance, which varies based on the shot size: Close shot (intimate), medium shot (social), or long shot (public/impersonal) (Kress & van Leeuwen, 2006).

c. Angle (involvement vs. detachment): The horizontal and vertical angle (e.g., frontal view for solidarity/involvement, oblique for detachment; high angle for power, low angle for submission).

2.2.3. Compositional Meaning

This function deals with how the visual elements are organized to create a coherent and salient message. The main components are as outlined in (4) below.

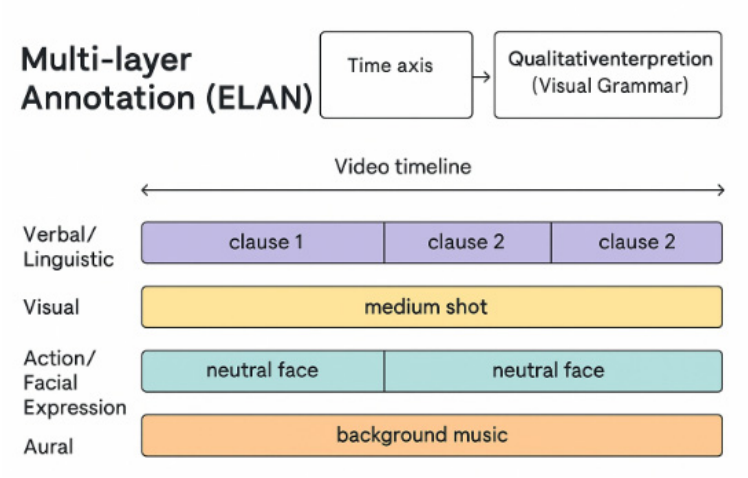
(4) a. Information value: Placement of elements, such as given/new (left/right) or ideal/real (top/bottom), to structure the message.

b. Salience: The degree to which an element draws attention through size, focus, contrast, or color (Dai & Jhang, 2025).

c. Framing: The presence or absence of connecting lines or borders that indicate whether elements belong together or are separate.

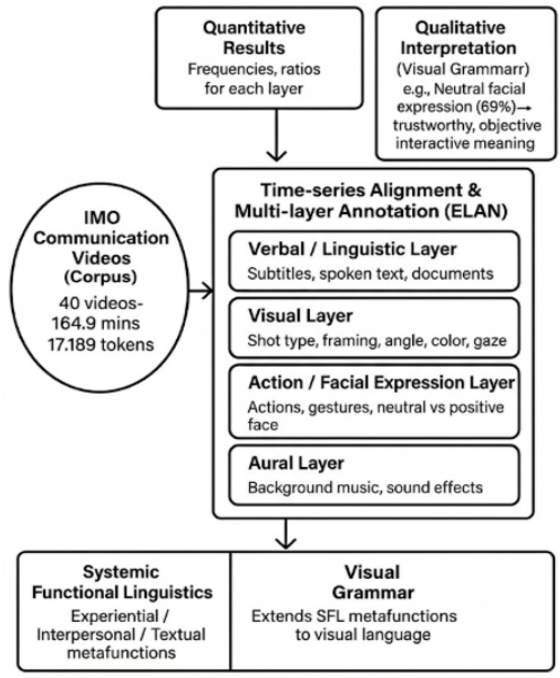
2.3. A Video Case Study: VG for Dynamic Discourse

The applicability of VG to highly temporal and multimodal communication is demonstrated by recent work such as Dai (2025), which conducts a corpus-based multimodal discourse analysis of the International Maritime Organization (IMO)'s communication videos on the marine environment. Drawing on time-aligned ELAN annotations, this study establishes a multi-layer system that captures the interaction between visual elements, action types, and emotional tones in video discourse. As summarized in Figure 1 below, integrating VG with time-aligned multimodal analysis provides a systematic way to trace how narrative structures, gaze behavior, and framing jointly construct environmental messages, thereby validating VG's potential for analyzing complex, concurrent visual communication such as sign language discourse.



[Figure 1] Time-Series VG Framework for Dynamic Video Analysis (Dai & Jhang, 2025)

To implement this validated approach, Figure 2 below displays the full integrated framework used in this study, demonstrating how the four distinct tracks of multimodal data are aligned for concurrent analysis across the three VG metafunctions.



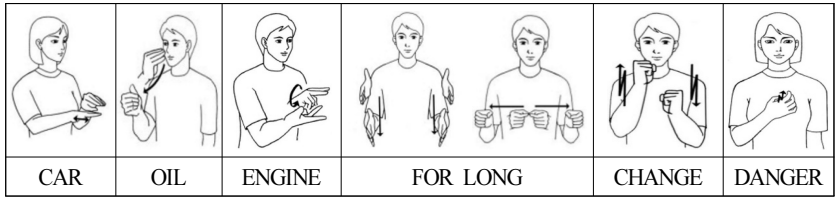
[Figure 2] Integrated Analysis Model of VG Framework for Dynamic Media Analysis (Dai & Jhang, 2025)

Figure 2 illustrates an integrated analysis model of the VG analytical framework applied to the IMO video corpus, detailing the specific components analyzed (e.g., Narrative Processes/Vectors, Gaze/Demand, Salience/Framing). Dai (2025) and Dai and Jhang (2025) provide a comprehensive integrated framework for analyzing the communication strategies in IMO videos. This model organizes the video content into four parallel, time-synchronized tracks: what is said, what is seen, the speaker's facial expressions/actions, and the audio/sound. This layered approach allows researchers to look at all features at once. By combining established theories like SFL and VG, the framework can both count what happens often (quantitative analysis) and explain how meaning is created (qualitative interpretation). Ultimately, it reveals how linguistic, visual, gestural, and sound elements work together to create the full meaning of the institutional message.

3. Data and Methodology

3.1. Data and Annotations

This study analyzes selected and transcribed examples of authentic discourse retrieved from the National Institute of Korean Language (NIKL) KSL Nuri Dictionary.[1] These examples were chosen for their clarity in demonstrating the simultaneous realization of manual and non-manual elements in various discourse contexts.[2] The qualitative analysis focuses on token occurrences of six focal lexical items (CAR, OIL, ENGINE, FOR LONG, CHANGE, and DANGER), which define the primary analytic dataset examined across the three metafunctions such as Rrepresentational, Interactive, and Compositional Meanings, as illustrated in Figure 3.



[Figure 3] Six Focal KSL Lexical Items Used in the Multimodal Analysis (KSL Dictionary, NIKL)

1) The KSL Nuri Dictionary (https://ksldict.korean.go.kr/kslserv), launched on February 6, 2025, is a new KSL–Korean dictionary developed by the NIKL. Diverging from the existing 'KSL Dictionary,' this new resource was compiled using corpus-based data to reflect authentic usage and linguistic conventions within the Deaf community.

2) The eleven video excerpts analyzed here are drawn from the KSL Nuri Dictionary, which comprises 1,131 lemmas, 287 variants, and 4,263 video recordings of authentic discourse. Excerpts were sampled using the following criteria: (i) clear visibility of both manual articulation and key non-manual elements (face and upper body), (ii) minimal occlusion and stable camera framing, and (iii) discourse contexts that allow the temporal coordination of manual, non-manual, and spatial–temporal resources to be observed. From these excerpts, the primary analytic dataset was defined as all token occurrences of six focal lexical items—CAR (24 tokens), OIL (2 tokens), ENGINE (2 tokens), FOR LONG (6 tokens), CHANGE (4 tokens), and DANGER (11 tokens)—yielding annotated tokens in ELAN. Materials outside these excerpts were consulted only for contextual understanding and are not included in the analytic counts reported in Section 4.

These six items, shown in Figure 3, were chosen because they effectively exemplify the complex interplay between manual elements, non-manual features, and spatial structuring necessary for demonstrating the validity of the VG framework. The detailed analysis of how these specific lexical items realize each of the three metafunctions is presented starting in Section 4.1.

The data were analyzed using the ELAN (EUDICO Linguistic Annotator; Wittenburg et al., 2006; Max Planck Institute for Psycholinguistics) software. ELAN's multi-tier annotation system is critical for capturing the simultaneity and temporal alignment of distinct semiotic modes (Lee & Jhang, 2025). Drawing on the established methodology used in multimodal video analysis (cf. Dai, 2025), the use of ELAN for sign language allows for precise, systematic timeline-based quantification of concurrent phenomena. ELAN is the gold standard for sign language transcription globally, with studies on multiple sign languages (e.g., American Sign Language (ASL), British Sign Language (BSL)) widely relying on its precision for aligning manual and non-manual data (Mondada, 2016).

3.2. ELAN Annotation Setup

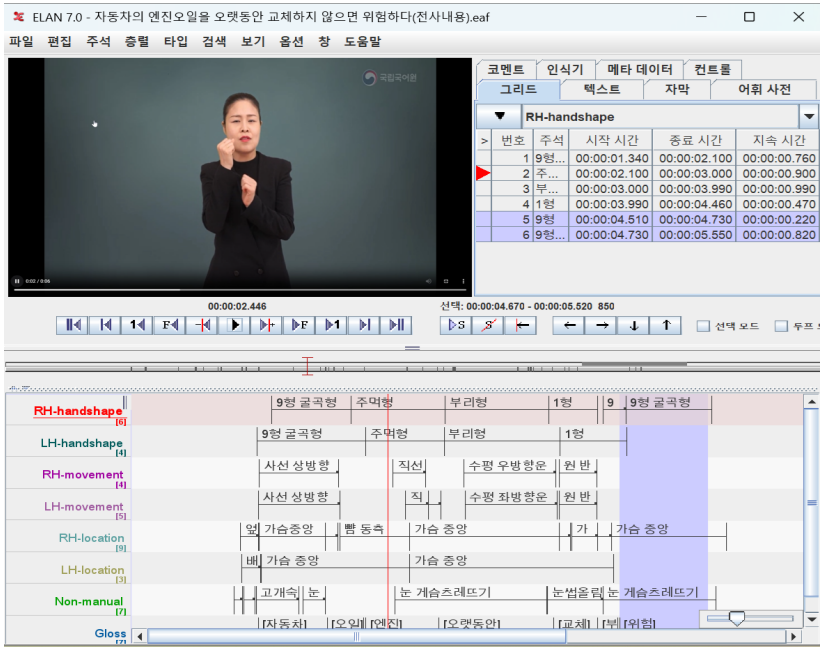
To systematically capture the concurrent features of KSL and operationalize the VG analysis, the video data were transcribed using ELAN software, which is optimized for multimodal analysis. The annotation structure was meticulously designed to create distinct, time-synchronized input tiers for all relevant semiotic channels.

The setup for KSL (Figure 4 below) was designed by adapting established multi-tier annotation models used in other sign language corpora, such as those developed for Hong Kong Sign Language (Fung et al., 2008). The final multi-tier structure provided the empirical foundation for separating and analyzing the semiotic contributions of each articulatory channel. The specific synchronized annotation tiers established for this study are detailed in (5).

(5) a. **Sign Gloss:** The lexical sign entry.
b. **Manual Features:** Handshape, Movement, Location.
c. **Non-Manual Features:** Facial Expressions, Gaze, Head/Body Posture.

d. **VG Metafunctions:** Dedicated tiers for annotating Representational, Interactive, and Compositional Meanings as they map onto the KSL elements, according to the conceptual framework in Table 1 (Section 3.3)

Figure 4 below illustrates the operational structure of this integrated ELAN–VG framework. This strategic setup ensures that the analysis can move beyond simple transcription, enabling both precise quantification and robust functional interpretation of meaning within dynamic KSL discourse.



[Figure 4] KSL ELAN Annotation Setup (Lee & Jhang, 2025)

3.3. Analytical Framework: Mapping VG to KSL

The core of this methodology is the systematic mapping of Kress and van Leeuwen's VG functions onto the inherent multimodal components of KSL (Dai & Jhang, 2025). The analytical framework systematically maps the three metafunctions of VG (Kress & van Leeuwen, 2020) onto the distinct articulatory

channels of KSL. This is necessary because KSL, as a dynamic, temporal, and multimodal language, requires adapting VG's concepts (originally designed for static imagery) to its linguistic structure. Our framework treats particular articulatory resources as tending to foreground different metafunctional work, while emphasizing semiotic simultaneity and multifunctionality; accordingly, Table 1 summarizes typical functional distributions and points of overlap rather than fixed one-to-one correspondences.

[Table 1] Functional Distribution and Overlap of Semiotic Resources in KSL within a VG Framework

Semiotic Resource	Representational Meaning	Interactive Meaning	Compositional Meaning
Handshape	Actor / Participant representation	Indexical stance (e.g., pointing, deictic engagement)	Salience through marked configuration
Movement	Process / Vector realization	Engagement through rhythm and intensity	Temporal framing and prominence
Location (Signing Space)	Circumstantial relations	Participant alignment	Information structuring (e.g., Given/New)
Facial Expression	Aspectual and affective meaning	Interpersonal stance (e.g., evaluation, emotion)	Emphasis and boundary marking
Gaze	Participant orientation	Viewer engagement / address	Framing and cohesion
Head / Body Posture	Event structure modulation	Attitudinal alignment	Discourse segmentation

Table 1 presents an analytically grounded account of KSL semiotic resources by highlighting their functional distribution and overlap within a VG framework. Rather than assuming fixed or exclusive correspondences between particular resources and metafunctions, the table foregrounds the principle of semiotic simultaneity, according to which meaning emerges from the coordinated interaction of multiple semiotic systems.

Within this framework, KSL resources are understood as inherently multifunctional. Non-manual elements such as facial expression and gaze may

extend beyond interactive functions to shape Representational Meaning, while manual elements including handshape and movement also contribute to Compositional Meaning through temporal and rhythmic structuring. Taken together, Table 1 should be understood as an exploratory analytical model that prioritizes functional affordances over mechanical reductionism and avoids a rigid division between manual and non-manual modalities. This non-reductionist mapping approach aligns with recent methodological developments in multimodal analysis, which emphasize functional overlap and inter-semiotic interaction over one-to-one correspondences between semiotic resources and analytical categories (Serengil & Ozpinar, 2024).

Guided by this framework, the following section presents a qualitative analysis of KSL discourse. Section 4 addresses the three research questions outlined in Section 1 by examining how representational, interactive, and compositional meanings are realized through the coordinated use of multiple semiotic resources.

4. Findings and Discussion

Section 4 presents a qualitative analysis of KSL discourse structured around the three research questions outlined in Section 1. Each subsection restates the core analytical focus of the corresponding research question in order to clarify how representational, interactive, and compositional meanings defined in VG are examined in the data.

As structured by the conceptual framework outlined in Table 1, this section presents the primary findings derived from the ELAN-annotated KSL data and illustrates how representational, interactive, and compositional meanings emerge through the coordinated deployment of manual, non-manual, and spatial–temporal resources. Rather than positing exclusive mappings, the analyses examine how particular resources often foreground specific metafunctional work while interacting with other channels in functionally overlapping ways. The lexical items in Figure 3 (CAR, OIL, ENGINE, FOR LONG, CHANGE, and DANGER) constitute the primary dataset analyzed across these three metafunctions.

4.1. Representational Meaning in KSL: Manual Elements as Actors, Processes, and Circumstances (Answering RQ 1)

Following the framework established in Section 3, this section addresses RQ1, which examines how Representational Meaning in VG is realized through the manual elements of KSL discourse. Focusing on handshape, movement, and location, the analysis investigates how these manual articulators function as semiotic resources for construing participants, processes, and circumstantial relations in KSL. The analysis demonstrates that Representational Meaning in KSL—the metafunction concerned with construing events, actions, participants, and their interrelations in the represented world—is systematically realized through the language's manual parameters of handshape, movement, and location.

In this analysis, the interpretation of handshape as realizing the role of an Actor and movement as realizing a Vector is grounded in the representational principles of VG. Within VG, Actors are defined as participants involved in depicted actions, while Vectors are understood as visually perceivable directional forces that establish processes, actions, and relations (Kress & van Leeuwen, 2006; 2020). Importantly, Vectors are not limited to graphic lines in static images, but encompass any form of directional movement that encodes actional meaning.

From this perspective, the application of the Actor–Vector distinction to KSL is not a metaphorical analogy but a theoretically motivated extension of VG to a visual–spatial linguistic modality. In KSL, handshape functions as a primary resource for identifying participants, whereas movement encodes the dynamic unfolding of actions through trajectory, directionality, and manner.

This finding provides initial qualitative support for RQ1 and suggests that the Actor–Vector distinctions in VG can be productively extended to KSL's dynamic visual–spatial modality. As illustrated in Figure 3, the lexical items CAR, OIL, ENGINE, FOR LONG, CHANGE, and DANGER provide a coherent dataset through which the semiotic contributions of these manual features can be examined in detail.

Figure 5 further visualizes the representational analysis of these items, demonstrating how each sign's handshape, movement, and location collectively construct the participant–process–circumstance configurations central to the

representational metafunction. Together, Figures 4 and 5 present the full set of representational resources that KSL draws upon to depict events, participants, and conditions within the discourse.



[Figure 5] ELAN-Based Representational Analysis of KSL

Handshape functions as the primary semiotic resource for encoding participants and their inherent physical or conceptual properties. For example, the 9-type curved handshape shared by the signs CAR and OIL iconically reflects salient characteristics of these referents—compactness, contour, or functional form—thus anchoring them as distinct participants in the represented event structure. In contrast, shifts to fist-shaped or beak-shaped configurations in ENGINE or FOR LONG signal changes in semantic class, degree of abstraction, or lexicalized function. These systematic alternations reveal that handshape operates not merely as a phonological unit but as a representational device capable of signaling taxonomic distinctions, scalar properties, and participant roles, a pattern clearly illustrated in Figure 5's handshape comparisons.

Movement, by contrast, constitutes the primary resource for encoding processes. Its function parallels VG's concept of the vector, which designates the trajectory or directionality that visually enacts an action, event, or change of state. In KSL, upward diagonal, horizontal, or circular repeated movements represent dynamic processes such as mechanical operation, transformation, or extended temporal duration. For example, the circular repeated movement in CHANGE conveys iterative or continuous action, while the extended linear movement in FOR LONG visually constructs the semantic domain of prolonged temporal span. These movement patterns create kinetic vectors that render the processual dimension of events highly visible, demonstrating that KSL naturally embodies the core principles of narrative representation identified in VG. Figure 5 depicts this mapping explicitly, showing how each sign's movement trajectory corresponds to its processual meaning.

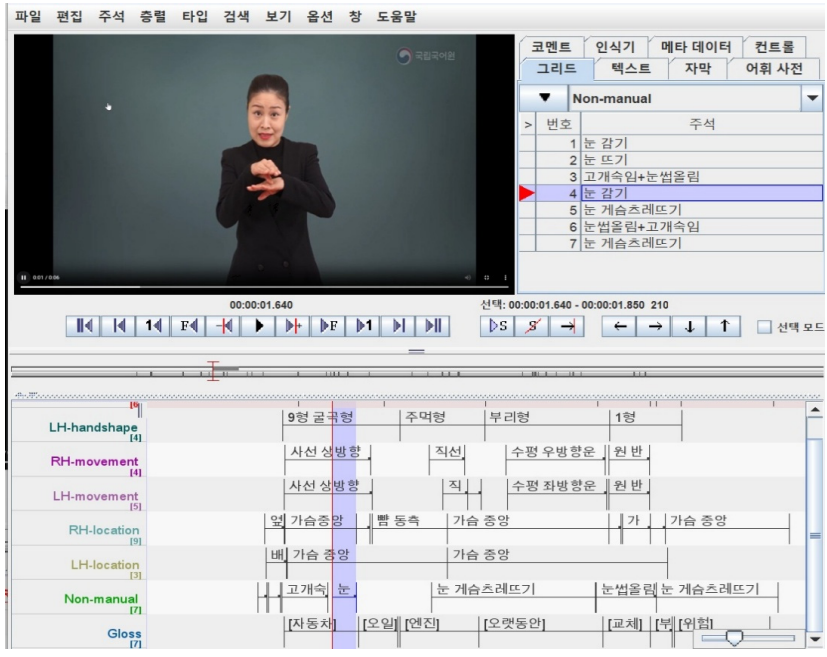
Location contributes circumstantial meaning by situating participants and processes within spatial, functional, or experiential domains of the signing space. Central chest articulation in DANGER, for instance, assigns the sign a high degree of affective and epistemic salience, marking it as a consequential or evaluative state. Conversely, lateral or lower-torso placements of CAR or ENGINE draw upon culturally shared embodied schemas—vehicles as external objects, engines as lower-body mechanical components—to anchor the discourse within familiar spatial frameworks. Such locational distinctions not only index spatial relationships but also encode subtle conceptual hierarchies, degrees of relevance, and pragmatic contrasts within the discourse. The spatial domains highlighted in Figure 5 show how signing-space organization reinforces circumstantial distinctions across lexical items.

Taken together, these manual parameters can be interpreted in terms broadly compatible with the representational categories articulated in VG. In this dataset, they tend to foreground participant, process, and circumstantial construal in ways that make the VG framework a plausible analytic lens for KSL discourse. Given the limited scope of the dataset, these observations are best understood as recurring patterns within the present data rather than definitive generalizations about KSL grammar.

4.2. Interactive Meaning in KSL: Non-Manual Elements as Attitude and Relation (Answering RQ 2)

Building on the representational analysis presented in Section 4.1, this section addresses RQ2, which explores how Interactive Meaning in VG is realized through non-manual elements in KSL. The analysis focuses on gaze, facial expression, and body posture, examining how these resources establish interpersonal relations, convey evaluative stance, and regulate engagement between the signer and the viewer.

The Interactive function in KSL, which centers on establishing and negotiating the social relationship between the signer and the viewer, is primarily carried by NMEs. These NMEs—including gaze direction, facial expressions, eyebrow movement, mouth morphemes, head orientation, and body posture—constitute the core interpersonal resources through which KSL manages engagement, stance, and interpersonal alignment.



[Figure 6] ELAN-Based Interactive Analysis of KSL

As visualized in Figure 6, the Interactive Meaning in KSL—the metafunction responsible for establishing, modulating, and negotiating the interpersonal relationship between the signer and the viewer—is realized primarily through NMEs. The ELAN-based representation in Figure 6 highlights the systematic layering of gaze direction, facial expression, eyebrow posture, mouth actions, and head–body orientation across the signing timeline, demonstrating that these NMEs operate as structured and non-random components of KSL discourse rather than as optional embellishments. Together, they form the core interpersonal resources through which KSL manages viewer engagement, epistemic stance, emotional tone, and evaluative positioning.

Within the VG framework, Interactive Meaning concerns the establishment of relations between represented participants and viewers through resources such as gaze, orientation, and expressive modulation (Kress & van Leeuwen, 2006). While NMEs in KSL—such as facial expression, gaze, and head or body posture—frequently foreground Interactive Meaning by signaling stance, evaluation, and address, their function is not restricted to the interpersonal domain.

Previous sign language research has shown that non-manual markers may also contribute to Representational Meaning by marking aspectual distinctions, event boundaries, or participant relations (Johnston, 2013). Accordingly, their analytical treatment in this study reflects functional overlap rather than categorical assignment.

Gaze functions as a fundamental structuring device for interpersonal alignment. As shown in the non-manual tier of Figure 6, direct gaze toward the viewer corresponds to what VG terms a Demand: it summons attention, signals emphasis, or invites a response, operating similarly to rhetorical questions or emphatic assertions in spoken languages. In contrast, gaze directed away, either toward the signing space or toward an imagined referent, performs an Offer, framing information as descriptive, neutral, or attributed to an external source rather than presented as a subjective evaluation. This distinction parallels the viewer–image relations in VG, where the presence or absence of direct gaze modulates interpersonal commitment.

Facial expressions contribute additional layers of interpersonal meaning by signaling attitude, evaluation, and affective stance. Figure 6 clearly illustrates the

alternation of eye closure, eyebrow raises, and widened or half-lidded eyes across different lexical items. These features serve as evaluative overlays that modify or intensify the meaning of the accompanying manual signs. For example, eye closure followed by opening functions as a discourse boundary marker, resetting attention before introducing a new informational unit. Raised eyebrows combined with a forward head tilt appear in evaluative or morally salient portions of the message—such as the segment involving CHANGE and NEGATION—thereby heightening the viewer's awareness of consequential or cautionary content. Meanwhile, half-closed eyes during signs such as FOR LONG or DANGER encode epistemic seriousness and temporal continuity, functioning analogously to the modality cues described in VG that soften, intensify, or recalibrate the interpersonal force of the message.

Head and Body Posture further regulates the communicative angle between signer and viewer. In VG, angle determines whether the represented participant establishes solidarity (front-on angle) or distance (oblique angle) with the viewer. Figure 6's ELAN alignment reveals that frontal body orientation coincides with assertions or key evaluative content, inviting the viewer into a shared perspective. In contrast, oblique angles or subtle head tilts mark moments where the signer transitions to a more descriptive mode, conveying detachment or depicting events from a more objective stance. These postural shifts contribute to the interpersonal positioning of the viewer, shaping how the message is interpreted and how authority, empathy, or evaluation is distributed across the discourse.

In summary, the non-manual behaviors detailed in the analysis—specifically gaze, facial expression, and head/body posture—play a central role in realizing Interactive (Interpersonal) Meaning as defined in VG, while also interacting with other channels and, in some contexts, contributing to representational and compositional work. These synchronized features are instrumental in regulating audience engagement, conveying emotional or epistemic stances, and directing the viewer's focus across the timeline. Consequently, in the present dataset, Interactive Meaning is frequently foregrounded through the coordination of non-manual resources, in interplay with manual articulation. This orchestration demonstrates that the interpersonal frameworks of VG are highly applicable to

the dynamic, embodied nature of sign language discourse. The analysis presented above is based on a restricted set of data and is therefore intended to highlight recurring tendencies rather than to advance comprehensive claims about the grammatical system of KSL.

4.3. Compositional Meaning in KSL: Spatial and Temporal Organization (Answering RQ 3)

Extending the discussion from the preceding analyses of Representational and Interactive Meanings, this section addresses RQ3, which investigates how Compositional Meaning in VG is constructed through the spatial and temporal organization of KSL discourse. Specifically, the analysis examines how the use of signing space and temporal patterning contributes to information structuring, salience, and discourse segmentation.

The spatial analysis presented in this section draws on the concept of information value as developed in VG. Kress and van Leeuwen (2006; 2020) formulated the Given/New distinction primarily with reference to two-dimensional visual layouts, such as images and page designs. The present study does not assume a direct equivalence between such planar structures and the three-dimensional, dynamically articulated signing space of KSL.

Instead, spatial contrasts in KSL are interpreted in functionally analogous terms. Signing space is organized relative to the signer's body as a semiotic center of reference, and prior research has demonstrated that spatial loci in sign languages are systematically employed to encode referential continuity, contrast, and discourse organization (Johnston, 2013). On this basis, left–right spatial distinctions are treated as discourse-organizational resources when they are deployed consistently within an utterance or across stretches of discourse.



[Figure 7] ELAN-Based Compositional Analysis of KSL

As illustrated in Figure 7, which presents the ELAN-based compositional analysis of the target utterance, KSL organizes meaning not only through individual signs but through the spatial and temporal structuring of signs across the signing space and along the discourse timeline. The figure highlights how manual and non-manual tiers are coordinated to create patterns of Information Value, Salience, and Framing, thereby offering a visual demonstration of how the compositional metafunction in VG is dynamically enacted in embodied sign language discourse.

The compositional function in KSL is realized through the strategic arrangement of signs within the physical Signing Space and through their distribution across the temporal axis of the utterance, as made visible in ELAN's multi-tiered timeline. Spatial placement primarily governs Information Value, determining which portions of the discourse are construed as Given, New, Ideal, or Real. Signs articulated in the high-salience region—such as the central chest or the dominant-hand frontal area—typically serve as focal elements carrying

major communicative weight. For instance, the sign DANGER occupies what VG identifies as the Centre/Real position, thereby elevating its pragmatic prominence and marking it as the most consequential outcome of the preceding conditions (CAR, OIL, ENGINE, FOR LONG, CHANGE/NEG). This spatial arrangement cues the viewer to interpret DANGER as the culmination of the narrative logic established earlier in the discourse.

Temporal organization is equally critical in shaping Compositional Meaning. ELAN's timeline visualization in Figure 7 reveals systematic patterns in movement speed, duration, and rhythmic contour that operate as cues for Salience. A lengthened movement or sustained hold phase draws heightened attention, functioning as a visual equivalent of rhetorical emphasis. For example, the extended movement during FOR LONG or the deliberate deceleration in the articulation of DANGER creates a temporal spotlight, signaling to the viewer that these segments are semantically and structurally pivotal. By contrast, sudden accelerations or intensified movement sequences foreground processes such as CHANGE or NEGATION, enhancing their perceptual and narrative prominence.

In addition to spatial and temporal structuring, KSL achieves Framing through shifts in non-manual markers, changes in rhythmic flow, and transitions between signing loci. Figure 7 captures several of these boundary-marking cues: fluctuations in eyebrow height, changes in eye aperture, and adjustments in head tilt frequently coincide with transitions between semantic units. These non-manual shifts serve a function analogous to framing devices in static visual media—such as borders, contrast, or spacing—that segment an image into meaningful parts while maintaining coherence across the whole. In KSL, such NMEs divide the discourse into thematic blocks, guiding the viewer through the internal structure of the message.

Ultimately, the compositional strategies observed in the discourse—centralized positioning for information value, rhythmic movement changes for salience, and boundary marking through non-manual shifts—illustrate how KSL can organize information through visual–spatial and temporal patterning in ways that are compatible with the compositional metafunction in VG. In the present dataset, these strategies suggest that information structuring in KSL can be usefully described in terms of information value, salience, and framing, while remaining

sensitive to context and to functional overlap with other semiotic channels. Accordingly, the spatial patterns discussed in this section should be understood as context-dependent configurations observed within the current dataset, not as universally applicable principles of KSL spatial grammar.

5. Conclusion and Implications

5.1. Summary of Findings

This study successfully applied Kress and van Leeuwen's VG to the analysis of KSL discourse, validating its feasibility and explanatory power for dynamic signed languages. The analysis identified patterned associations between KSL's multimodal resources and VG's metafunctions. In the present dataset, manual articulators (handshape, movement, location) often foreground representational construal (participants, processes, circumstances), while non-manual features (e.g., gaze and facial expression) frequently foreground interactive work (engagement, stance, evaluative positioning). Spatial–temporal organization contributes centrally to Compositional Meaning (information structuring, salience, segmentation), with meaning emerging through coordinated use and functional overlap across channels rather than through exclusive one-to-one mappings.

5.2. Theoretical Contribution

The primary theoretical contribution is the successful extension of VG from static visual media to a fully dynamic, spatially and temporally structured linguistic system—sign language. By applying VG's metafunctional lens to sign language discourse and examining how meanings are distributed across coordinated articulatory resources, this research offers initial qualitative support for the plausibility of extending metafunctional analysis beyond static visual media. More cautiously, the findings are consistent with the social semiotic view that metafunctional principles may be interpretable across modes, while underscoring the need for broader empirical testing to determine the extent and limits of such cross-modal generalizability.

5.3. Implications and Future Research

While this study successfully demonstrates the theoretical extension and analytical potential of the VG framework, it is grounded in a qualitative analysis of a limited corpus of KSL discourse. Therefore, the identified patterns, while systematic, should be viewed as preliminary findings that require further statistical validation for broader generalization.

While limited in scope, the present findings demonstrate the potential of VG as a theoretically informed framework for analyzing KSL discourse. The patterns identified here should therefore be understood as preliminary and hypothesis-generating, inviting further investigation using larger and more diverse sign language corpora.

This research nonetheless carries meaningful practical implications for three key areas. In sign language education, the framework underscores the central role of NMEs in constructing Interactive and Compositional Meanings, offering educators a structured way to teach how gaze behavior and facial expressions contribute to discourse interpretation. In professional interpreting, the framework provides a principled method for identifying discourse intent—particularly evaluative stance and affective nuance—which may help interpreters improve accuracy and consistency. In the context of sign-language media, the model offers an effective analytical tool for understanding how spatial–temporal structuring and non-manual emphasis are employed rhetorically to direct viewer attention and guide interpretive outcomes.

Moving forward, further research is needed to deepen and extend these findings. First, large-scale quantitative validation using more extensive KSL corpora should be conducted to test the statistical robustness of the identified patterns and refine the analytical categories. Second, cross-linguistic comparative studies involving other sign languages such as ASL, BSL, or JSL are recommended to examine the universality of VG's metafunctions within signed modalities and to identify language-specific or culture-specific adaptations. Ultimately, such cross-modal and cross-linguistic studies will not only enhance our understanding of sign language structure but also clarify how metafunctional organization is realized differently across communicative modes.

References

Anderson, C., Bjorkman, B., Denis, D., Doner, J., Grant, M., Sanders, N., & Taniguchi, A. (2022). *Essentials of linguistics* (2nd edition.). Toronto, Ontario: eCampusOntario.

Dai, Z. (2025). *A corpus-based multimodal discourse analysis of IMO's communications on the marine environment* (Unpublished Doctoral Dissertation). Korea Maritime & Ocean University.

Dai, Z. & Jhang, S. (2025). A corpus-based multimodal discourse analysis of IMO's marine environment communication videos using Visual Grammar. Paper presented at the The Mirae English Language and Literature Association 2025 Winter Conference on December 13, Jeju National University, Jeju, South Korea.

Fung, C., Lam, S., Mak, J., & Tang, G. (2008). *Establishment of a corpus of Hong Kong Sign Language acquisition data: From ELAN to CLAN*. Centre for Sign Linguistics and Deaf Studies, The Chinese University of Hong Kong.

Halliday, M. A. K. (1985). *An introduction to functional grammar* (1st edition). London: Edward Arnold.

Johnston, T. (2013). Towards a comparative semiotics of pointing actions in signed and spoken languages. *Gesture, 13*(2), 109-142.

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

Kress, G. & van Leeuwen, T. (2006). *Reading images: The grammar of visual design.* (2nd Edition). London: Routledge.

Kress, G. & van Leeuwen, T. (2020). *Reading images: The grammar of visual design.* (3rd Edition). London: Routledge.

Lee, S. & Jhang, S. (2025). A preliminary study of multimodal analysis of Korean Sign Language through Visual Grammar. Paper presented at the 4th Interdisciplinary Research Workshop on Sign Language, Busan, South Korea.

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics, 20*(3), 336-366.

National Institute of Korean Language. (n.d.). *Korean Sign Language Dictionary*. https://sldict.korean.go.kr

Serengil, S. & Ozpinar, A. (2024). A benchmark of facial recognition pipelines and co-usability performances of modules. *Journal of Information Technologies, 17*(2), 95-107.

Wittenburg, P. Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation* (LREC'06), Genoa, Italy. European Language Resources Association (ELRA).

Lee, Seonhye
45 Yongso-ro, Nam-gu, Busan 48513, South Korea
Lecturer
Pukyong National University
Phone: 051-629-5370
Email: shlee2016a@naver.com

Dai, Ziyun
727 Taejong-ro, Yeongdo-Gu, Busan 49112, South Korea
PhD student
Korea Maritime & Ocean University
Phone: 051-410-4590
Email: ameliadai1017@naver.com

Jhang, Se-Eun
727 Taejong-ro, Yeongdo-Gu, Busan 49112, South Korea
Professor
Korea Maritime & Ocean University
Phone: 051-410-4590
Email: sejhang@naver.com

Received on December 17, 2025
Revised version received on December 26, 2025
Accepted on December 31, 2025