

The Player's Multimodal Encounter: Character Image Construction and Narrative Formation in *Arknights*

Mingyu Deng

School of Applied Foreign Language, Guangzhou Huali College

Author Note

The author declares no conflicts of interest to disclose. Correspondence concerning this article should be addressed to Mingyu Deng, School of Applied Foreign Language, Guangzhou Huali College, Guangzhou, Guangdong, China. Email: sylvantetsuya@foxmail.com

Abstract

Digital games increasingly ask players to recognize a character before that character has been fully explained by plot. A mobile role-playing game such as *Arknights* does not build characterization through verbal lore alone; it coordinates costume, posture, gaze, color, interface position, voice lines and background writing into a dense multimodal field. Existing linguistic research on games has often emphasized localization, player interaction, or literacy learning, while game studies has tended to discuss narrative and mechanics as medium-specific forms. This paper argues that an applied-linguistic account of digital games must attend to a distinct mode of meaning-making: the construction of characters and story worlds through recurrent visual-cultural cues that are immediately legible to players. This paper introduces the concept of “multimodal prototype analysis” - semiotic residues of myth, art history, folklore, professional discourse, and popular genre conventions that are embedded in game character images and activated through visual grammar. Unlike explicit background stories, which tell players what a character has done, multimodal prototypes guide players toward what kind of character they are encountering before the story is fully narrated. Drawing on social semiotics, Kress and van Leeuwen's visual grammar, systemic functional linguistics, and game narrative theory, the paper proposes “prototype resonance” as a process through which players infer role, morality, affiliation, and narrative tension from the interaction between visual design and textual framing. The empirical portion reconstructs and rewrites a case study of *Arknights* based on a corpus of five interface images and fifty characters images. Findings suggest that character images in the corpus are dominated by symbolic representation, demand-oriented gaze, socially distant framing, and center-margin composition. These features do not merely decorate characters; they convert cultural prototypes into playable narrative expectations and transform routine interface operations into world-internal actions. The paper concludes by outlining a critical multimodal game literacy for analyzing how digital games organize narrative participation through character design.

Keywords: multimodal discourse analysis, visual grammar, *Arknights*, game narrative, character image

Received 15 July 2026; revised 23 August 2026; accepted 13 September 2026; available online 25 September 2026; version of record 30 September 2026.

Citation: Deng, M. Y. (2026). The player's multimodal encounter: Character image construction and narrative formation in *Arknights*. *Journal of Language*, 2(3), 88–112. <https://doi.org/10.64699/26AJFV6045>

1. The Player's Multimodal Encounter: A Social-Semiotic Inquiry into Character Image and Game Narrative

Contemporary video games rarely introduce characters in a purely verbal way. Before a player reads a complete biography, watches a plot sequence, or finishes a chapter, the game has already offered a portrait: a body positioned in space, a gaze directed toward or away from the player, a weapon, a color palette, a name, a costume, a background emblem, and a set of interface affordances that tell the player how this figure should be handled. In this sense, the playable character arrives as a multimodal event before becoming a narrated person. The player is invited to interpret the character not only by reading but by recognizing patterns in images, spatial layouts, and culturally familiar signs.

This paper examines that interpretive process through the case of *Arknights*, a mobile tactical role-playing game whose operators are built from combinations of visual design, mythological reference, professional role, and short textual lore. The central question is how videogame character images construct both character identity and game narrative through multimodal resources.

The argument is grounded in applied linguistics because digital games are not only entertainment systems but communicative environments. They organize address, participation, literacy, stance, narrative inference, and cross-cultural meaning-making. As Gee (2003) argues, games can be understood as semiotic domains in which players learn how signs, practices, identities, and values are connected. The New London Group's (1996) account of multiliteracies similarly asks researchers to attend to meaning across linguistic, visual, audio, spatial, and gestural modes. A game interface is therefore not a neutral technical surface. It is a designed discourse that tells the player who they are, what counts as action, and how the fictional world should be read.

Arknights is a useful case precisely because it makes this discourse visible. Its setting

places the player in the role of the “Doctor,” a figure who commands operators, reads mission data, enters combat stages, manages resources, and encounters characters through files, voice lines, and visual portraits. Many character designs allude to recognizable cultural prototypes: Surtr evokes the Norse fire giant; Warfarin combines vampire iconography with medical terminology; Nightingale layers the nurse Florence Nightingale with the image of a caged bird; Phantom draws on the gothic and operatic figure of The Phantom of the Opera; Pallas activates classical Greek imagery. These allusions are interpretive shortcuts that make fictional characters narratively graspable.

These shortcuts are called “multimodal prototype traces.” A prototype trace is different from a citation, a quotation, or a simple reference. It is a compressed semiotic residue that survives across modes: a color scheme, a pose, a prop, a name, a genre convention, a mythic creature, or a professional symbol. When these traces are placed in a game portrait and reinforced by short pieces of lore, they allow the player to infer character identity before receiving a full story. The trace is both visual and narrative. It points backward to a cultural memory and forward to a playable expectation.

The paper makes three claims. First, character images in games can be analyzed as multimodal texts whose representational, interactive, and compositional meanings participate directly in narrative construction. Second, the character design of *Arknights* relies heavily on prototype traces that transform older cultural materials into new gameplay identities. Third, the interface of the game turns everyday operations such as logging in, shopping, selecting friends, and entering combat into diegetic actions, thereby linking player behavior to the narrative role of the “Doctor.”

The paper proceeds in five sections. Section 2 develops the theoretical framework by bringing together social semiotics, visual grammar, applied-linguistic approaches to games, and game narrative theory. Section 3 presents the corpus and method, then analyzes interface

images and character images from *Arknights*. Section 4 discusses the broader implications of prototype resonance for multimodal discourse analysis and game literacy. Section 5 concludes by identifying the contributions and limitations of the study.

2. Theoretical Framework

2.1 Digital Games as Multimodal Discourse

Multimodal discourse analysis begins from the premise that language is one semiotic resource among others. Halliday's social semiotic view of language emphasizes that meaning is functional, situated, and socially organized (Halliday, 1978; Halliday & Matthiessen, 2014). Kress (2003, 2010) extends this orientation to contemporary communication by arguing that different modes have different affordances: image, writing, speech, gesture, layout, color, sound, and movement do not simply repeat one another. They distribute meaning across a communicative ensemble.

Digital games intensify this distribution. A player may read a line of dialogue, hear a voice, inspect a portrait, compare equipment icons, respond to a map, and press a button in the same moment of play. The meaning of the moment is not located in any one mode alone. It emerges from orchestration. O'Halloran (2004) and Jewitt et al. (2016) emphasize that multimodal analysis must identify how modes cooperate, compete, and specialize. In games, this cooperation is also procedural: the interface does not only show meanings but lets the player act on them.

This matters for applied linguistics because games are spaces of literacy practice. Gee (2003) treats games as environments in which players learn to read signs in relation to situated identities and social practices. Thorne et al. (2009) show that online gaming can become a site of language socialization and intercultural participation. Ensslin (2011) further demonstrates that gaming has a language of its own, shaped by player address, genre conventions, instruction, and gameplay interaction. The question, then, is not whether games are relevant to language

study. The question is how precisely their multimodal signs produce discourse effects.

The present paper focuses on two discourse effects: character image construction and narrative construction. Character image construction refers to the semiotic work through which a fictional figure is made legible as a type of person, agent, professional role, moral position, or social relation. Narrative construction refers to the arrangement of visual, textual, spatial, and procedural cues into a sense of story-world, role, conflict, and temporal development. In *Arknights* these effects are inseparable: the character image is a narrative device, and the interface is a character-addressing device.

2.2 Visual Grammar and the Three Metafunctions

Kress and van Leeuwen's visual grammar is especially useful because it translates the systemic-functional concern with metafunctions into visual analysis. Their framework distinguishes representational meaning, interactive meaning, and compositional meaning (Kress & van Leeuwen, 2021). Representational meaning concerns what is depicted and how participants, actions, concepts, and symbolic attributes are arranged. Interactive meaning concerns the relation between the depicted participant and the viewer, including gaze, social distance, angle, and modality. Compositional meaning concerns information value, salience, framing, and the distribution of elements across the visual field.

These three dimensions work together. A character who looks directly at the viewer, stands at full-body distance, occupies the center of the image, and holds a glowing staff offers a different interpersonal and narrative invitation from a character who turns away, appears partially obscured, and is framed by ruins or shadow. The first may be read as accessible, allied, or summonable; the second as distant, mysterious, or narratively unresolved. Visual grammar gives the analyst vocabulary for explaining such differences without reducing them to simple description.

In game character design, representational meaning often appears through symbolic

processes. A symbolic process presents a participant as the bearer of attributes. A sword, shield, cage, mask, medical bag, musical instrument, or animal feature can tell the player what role the character occupies. Interactive meaning then positions the player in relation to that role. A demand image uses gaze to establish a relation with the viewer; an offer image presents the character as an object of contemplation or narrative distance. Compositional meaning makes some features central and others marginal, thereby guiding attention and interpretation.

The central analytic move in this paper is to treat visual grammar not only as a description of static images but as a bridge between image and play. A centered character portrait does not simply have high salience. It becomes the visual anchor through which the player decides whether to recruit, command, trust, or investigate a character. A demand gaze does not simply create eye contact. It helps naturalize the player's role as the addressed Doctor. A symbolic attribute becomes a narrative key.

2.3 Game Narrative as Designed Semiotic Environment

Game studies have long debated the relation between games and narrative. Aarseth (1997) emphasized the ergodic nature of cybertexts, in which users must perform nontrivial work to traverse the text. Juul (2005) described video games as "half-real," combining real rules with fictional worlds. Salen and Zimmerman (2004) likewise frame games as systems of rules, play, and culture. These accounts caution against treating games as novels with buttons attached.

At the same time, narrative remains central to many games. Jenkins (2004) proposes "narrative architecture" as a way of understanding how game spaces can evoke, enact, embed, and generate story. Ryan (2006) similarly argues that digital media expand narrative by introducing interactivity, spatial exploration, and multiple forms of immersion. The question is not whether games are narrative or non-narrative, but how different semiotic systems make narrative possible under conditions of play.

Arknights is especially relevant to this question because it tells story through fragments. The player encounters main plot chapters, operator files, event stories, combat maps, interface slogans, recruitment materials, and voice lines. A character may be introduced through a name and portrait before the player reads the complete file. The narrative therefore depends on anticipatory interpretation. The player is asked to infer the possible story from the available signs.

This anticipatory structure is where multimodal prototype traces become important. A prototype trace allows a game to compress a large cultural narrative into a small sign. A mask and an opera-house palette can call up the gothic romance of *The Phantom of the Opera*. A red sword and name can call up *Surtr* and the destructive fire of Norse myth. A white medical figure named *Nightingale* can call up nursing, care, war trauma, and sacrifice. These traces make narrative economical. They do not replace game lore, but they prepare the player to read it.

2.4 Defining Multimodal Prototype Traces

The concept of a multimodal prototype trace names a particular category of game sign. It is a recognizable residue of a prior cultural form, carried across one or more modes and placed inside a new game context. It differs from explicit lore because it is not primarily propositional. It does not say, “this character is protective” or “this character is tragic.” Rather, it arranges cues that make such inferences available: a shield, a frontal stance, a direct gaze, a guardian-animal motif, a subdued palette, or a textual background that confirms the visual premise.

Prototype traces also differ from simple intertextual references. Intertextuality can involve quotation, parody, adaptation, allusion, or genre memory (Fairclough, 1992). A prototype trace is more operational. It is designed to work quickly in a playable environment. The player may not identify the exact source, but the trace can still organize interpretation. A

player who does not know Komainu by name may still read Hung as a guardian because the shield, animal features, posture, and moral framing converge. A player who has never read Dracula may still read Warfarin through vampire conventions because pale skin, red eyes, blood imagery, and medical discourse intersect.

Table 1

Explicit Narrative Exposition Versus Multimodal Prototype Traces

Dimension	Explicit Narrative Exposition	Multimodal Prototype Traces
Primary modality	Verbal lore, dialogue, operator files	Image, name, prop, color, layout, voice line, interface position
Type of information	Propositional and sequential	Associative, symbolic, and culturally sedimented
Typical question answered	What happened to this character?	What kind of character is this?
Temporal effect	Builds story across time	Compresses story before full exposition
Player activity	Reading, remembering, connecting plot facts	Recognizing, inferring, aligning with prior cultural knowledge
Narrative function	Expands background and motivation	Prefigures role, stance, mood, and expectation

2.5 Prototype Resonance and the Gameplay Character

When a player recognizes a prototype trace, the result is not merely identification of a source. The more important process is the “prototype resonance.” Prototype resonance occurs when visual-cultural clues activate prior knowledge strongly enough to shape the player’s understanding of a new character. The player sees the trace, connects it to what they already know, and uses that knowledge to interpret the game character.

Prototype resonance is sometimes unconscious. Some players may explicitly identify Pallas with Athena or Phantom with The Phantom of the Opera. Others may only feel that Pallas is classical, ritual, heroic, and priestly, or that Phantom is gothic, theatrical, and distant. This distinction matters. A prototype trace can work at different levels of cultural literacy. It

can reward expert recognition while still producing a general mood or role for less specialized players.

The “gameplay character” is therefore a figure constructed through both narrative and play. It is a playable sign-complex: a visual portrait, a set of skills, a voice, a role in combat, a relation to the player, and a possible object of collection or investment. The gameplay character is built when these semiotic resources cohere. If the visual image, background text, combat role, and interface behavior contradict one another without design purpose, the character may feel arbitrary. If they resonate, the player experiences a sense of narrative coherence before the story has been completely explained.

2.6 Synthesis: From Visual Grammar to Character-Narrative Coupling

The framework can now be summarized in three linked propositions. First, visual grammar identifies how character images generate representational, interactive, and compositional meanings. Second, multimodal prototype traces explain why certain visual details carry narrative force: they are sedimented cultural signs that connect a new character to older repertoires. Third, prototype analysis describes the player’s interpretive uptake of these signs. Together, these propositions describe “character-narrative coupling,” the process by which games make character design carry the burden of story-world construction.

This framework generates the empirical question that guides the remainder of the paper: how do the interface images and character images of *Arknights* coordinate visual grammar, cultural prototypes, and textual lore to construct characters and narrative participation?

3. Empirical Study

3.1 Corpus and Analytical Unit

The empirical material corpus derives from five frequently encountered interface images from *Arknights* - the store interface, login interface, friend interaction system, main operating interface, and battle interface - and fifty characters images selected from official or

game-internal sources. Among the character images, ten were chosen for close qualitative interpretation: Shalem, Liskarm, Phantom, Surtr, Nightingale, Warfarin, Czerny, Hung, Skadi, and Pallas.

The analytical unit is an image-text complex. For interface examples, this includes screen layout, character address, functional buttons, color salience, and the player's implied action. For character examples, this includes the operator portrait, name, posture, props, background elements, gaze, color palette, and relevant lore summaries. Because the purpose of this rewritten article is academic analysis rather than visual reproduction, copyrighted game images are described rather than redistributed.

The study is qualitative with descriptive quantitative support. The quantitative component comes from coding of fifty characters images according to visual grammar categories. These counts are used as descriptive indicators of patterning, not as inferential statistics. The qualitative component consists of close multimodal readings that interpret how selected images coordinate with textual lore to construct character and narrative.

3.2 Methodology: A Three-Stage Multimodal Coding Process

The analysis follows a three-stage process adapted from multimodal discourse analysis and qualitative game analysis (Consalvo & Dutton, 2006; Kress & van Leeuwen, 2021). The purpose is aiming to reconstruct the semiotic invitations built into the game materials.

3.2.1 Stage 1: Visual Grammar Coding

In the first stage, each character image was coded for the three meanings of visual grammar. Representational meaning was coded by identifying whether the image primarily involved narrative action or conceptual-symbolic attribution. Interactive meaning was coded through visual contact, social distance, perspective, and modality. Compositional meaning was coded through information value, salience, and framing.

3.2.2 Stage 2: Prototype Trace Identification

In the second stage, visual and textual details were read for cultural prototypes. A prototype trace was coded when a name, prop, posture, species feature, color arrangement, professional symbol, or background motif evoked a recognizable cultural source or type. The source could be mythological, historical, literary, professional, or genre-based. The analysis then asked how the trace was transformed in the game context. For example, the vampire trace in Warfarin is not simply reproduced; it is redirected into a medical and ethical role. The point is the explanation of how older signs are made playable.

3.2.3 Stage 3: Character-Narrative Coupling

In the third stage, the study examined how image, text, and interface action worked together. This stage asked three questions. First, what character identity is proposed by the visual and textual ensemble? Second, how is the player positioned in relation to that identity? Third, what aspect of the wider game narrative is made intelligible through the character or interface? The analysis therefore moves from visual feature to discourse function.

3.3 Findings

The analysis revealed four primary findings. First, the interface of *Arknights* constructs operational narration: routine player actions are framed as actions performed by the Doctor inside the fictional world. Second, character images are dominated by symbolic representation, which makes them especially suited to prototype-based characterization. Third, demand gaze and center-margin composition establish a patterned relation between character and player. Fourth, cultural prototype traces allow the game to compress complex narrative identities into immediately legible images.

3.3.1 Interface Images as Operational Narration

The five interface examples show that *Arknights* narrativizes ordinary game operations. The store interface does not merely display purchasable items. It presents a shop assistant

whose gaze and posture simulate a salesperson-customer relation, turning economic exchange into an interpersonal encounter. The login interface does not merely provide access. Its wording about establishing a connection with Rhodes Island turns technical entry into a diegetic act of joining the organization. The friend interaction system does not offer open chat, but it makes social exchange visible through credit points and base visits. The main interface positions the selected assistant operator as a subordinate or companion who addresses the Doctor, often through short lines that simulate workplace interaction. The battle interface places completed stages and target stages in a spatial sequence, transforming level selection into tactical progression.

Table 2

Interface Screens and Their Narrative Functions

Interface screen	Dominant multimodal cue	Player position constructed	Narrative function
Store interface	Shop assistant gaze, product-centered layout, organizational background	Customer and institutional member	Converts purchase into interpersonal service inside the game world
Login interface	Connection wording, abstract technical screen, center-margin layout	Returning Doctor reconnecting to Rhodes Island	Converts app entry into re-entry into the fictional organization
Friend interaction system	List layout, access buttons, credit exchange	Socially networked player-operator	Converts resource exchange into mediated social participation
Main operating interface	Assistant operator, work-panel divisions, direct address	Leader receiving reports and assistance	Converts menu navigation into workplace command
Battle interface	Given-new progression, marked target stage, red salience	Tactical commander selecting the next conflict	Converts level selection into military decision-making

Note. The interface is the main mechanisms by which the player's routine actions are made narratively meaningful.

These interface findings matter because they show that narrative construction is procedural as well as visual. The interface repeatedly addresses them as someone who logs into Rhodes Island, buys supplies, manages contacts, gives commands, and selects missions. This

is an applied-linguistic phenomenon because identity is produced through address and action: the second-person position of the player is constructed by a recurring multimodal discourse.

3.3.2 Symbolic Representation and the Dominance of Prototype Traces

The character-image coding shows a strong preference for symbolic representation. Of the fifty character images summarized, forty-one were coded as conceptual or symbolic processes, while nine were coded as narrative processes. This distribution suggests that *Arknights* character portraits usually present characters as bearers of attributes rather than as participants in unfolding actions. That design choice is narratively efficient. A portrait is expected to reveal who the character is, not merely what the character is doing.

Table 3

Distribution of Visual Grammar Features in Fifty Character Images

Visual grammar dimension	Dominant category	Count	Percent age	Narrative implication
Representational meaning	Conceptual/symbolic process	41/50	82%	Characters are introduced through attributes, symbols, and cultural prototypes.
Representational meaning	Narrative process	9/50	18%	Action is present but less central than identity-bearing design.
Visual contact	Demand	45/50	90%	Most operators visually address the player and support the Doctor relation.
Visual contact	Offer	5/50	10%	A smaller group is presented as distant, elusive, or narratively withheld.
Social distance	Socially distant	50/50	100%	Full-body or near full-body portraits frame characters as inspectable units of role and equipment.
Information value	Center-margin	46/50	92%	The character body is the interpretive center; background and symbols orbit it.
Information value	Given-new	4/50	8%	Rare exceptions preserve a source composition or emphasize contrastive development.

Note. Percentages are calculated from the counts preserved in the image corpus. They are descriptive, not inferential.

The dominance of symbolic processes is significant. A symbolic image invites the viewer to read attributes, not events. In *Arknights*, those attributes are rarely invented from nothing. They are assembled from cultural repertoires: mythic beings, professional icons, literary figures, art-historical poses, genre colors, and recognizable tools. The character portrait becomes a place where old signs are recombined into new identities.

This design pattern also explains why character images can carry narrative information before plot exposition. A player who sees Warfarin's blood-red palette, bat-like imagery, and medical role can infer a tension between danger and care. A player who sees Nightingale's lamp-like staff, fragile posture, and caged-bird imagery can infer healing, trauma, and confinement. A player who sees Pallas framed by Greek ritual imagery can infer priesthood, heroic memory, and a relation to classical antiquity. These inferences are guided by prototype traces.

3.3.3 Demand Gaze, Social Distance, and the Player's Role

The corpus also shows a patterned relation between gaze and player address. Approximately nine out of ten character images were coded as demand images. In Kress and van Leeuwen's (2021) terms, demand images create a relation between the represented participant and the viewer through gaze. In a game context, this relation has a special function: it helps establish the player as an addressee within the fictional world.

The player of *Arknights* is not positioned as a detached spectator. The player is addressed as Doctor, a role that combines command, care, memory, and responsibility. Demand gaze supports this structure by making operators appear available to the player as interlocutors, allies, or subordinates. Even when the portrait is socially distant, the direct gaze can create interpersonal address. The player sees the full body, equipment, and background symbols, but also receives the character's attention.

The combination of demand gaze and social distance may appear contradictory at first.

Direct gaze suggests intimacy, while full-body distance suggests inspection. In the game context, however, the combination is productive. The operator is personally available but operationally assessable. The player is looking at a person; and the player is also evaluating a recruit, teammate, or narrative figure. This double relation is central to the gameplay character. The character must be emotionally legible and functionally usable.

Offer images are therefore especially meaningful when they occur. Phantom, for example, is read through distance and withholding. His visual orientation and gothic prototype do not invite the same direct interpersonal bond as many allied operators. The offer structure supports his narrative aura: he is present as an object of fascination, rumor, and danger rather than as a fully transparent companion. In this case, the exception confirms the rule. When the demand pattern is dominant, the offer pattern becomes a marked sign of narrative distance.

3.3.4 Center-Margin Composition and Narrative Salience

Almost all character images (92%) in the corpus use a center-margin information structure. The character body occupies the center, while background motifs, props, symbolic animals, light effects, and environmental details appear around it. This composition performs two functions at once. It makes the character commercially and visually salient, and it organizes background information as interpretive support.

Center-margin composition is especially important in a character-collection game. A portrait must be quickly recognizable. It must tell the player where to look first. Yet the margin is not empty. It often contains the prototype trace that deepens characterization: a piano behind Czerny, a sea-like blue field around Skadi, a cage or bird around Nightingale, blood-red effects around Warfarin, classical ritual signs around Pallas, or a shield and guardian-animal coding around Hung. The center answers “who is this?” while the margin answers “what world of meaning surrounds this figure?”

There is an instructive given-new exception reported in the corpus because it suggests

a different type of visual argument. Given-new composition can preserve the memory of an art-historical source or create contrast between known and emerging meanings. In Liskarm's case, the protective stance, shield, and upward visual energy can be read against older revolutionary or guardian imagery. The design uses composition to make a familiar value - protection, liberty, defense - available in a new military-technological context.

3.3.5 Character Case Readings: Prototype Resonance in Practice

The ten close readings show how prototype resonance operates across different cultural sources.

Shalem is framed through an art-historical and theatrical sensibility. The image identifies an association with Bacchic imagery and a deep relation to beauty, desire, and destructive aesthetic passion. The character's posture, cane-like prop, and background story produce a figure whose identity is not reducible to combat function. The prototype trace here is aesthetic excess: the character is read through art, performance, longing, and moral instability.

Liskarm is constructed through protection and civic force. Her shield is more than equipment. It is a symbolic attribute that turns the body into a defensive sign. The image associates her composition with liberty and revolutionary iconography, while her lore emphasizes a desire to bring stability rather than chaos. Whether or not every player identifies the precise source, the prototype resonance is clear: Liskarm is a protector whose legitimacy comes from disciplined force used for others.

Phantom offers a particularly clear example of literary and genre transfer. His name, mask, cloak, theatrical setting, and French-inflected reference evoke *The Phantom of the Opera*. The image is not simply "about" the musical. It translates gothic secrecy, damaged artistry, and unstable desire into a playable operator identity. The fact that Phantom is one of

the more distant or offer-like figures in the corpus reinforces this narrative position. He is seen, but not easily possessed by the viewer's understanding.

Surtr demonstrates mythological compression. In Norse mythology, Surtr is associated with fire, destruction, and a flaming sword. In *Arknights*, this prototype is transformed into a young female operator whose red palette and huge sword preserve the destructive fire trace while altering gender, genre, and commercial appeal. The result is neither direct adaptation nor arbitrary borrowing. It is gameplay reconfiguration: mythic force is made playable through character design.

Nightingale is built through layered metaphor. The name first evokes Florence Nightingale and the role of medical care. The lamp-like or glowing staff strengthens the healing prototype. At the same time, the image of the nightingale as bird, sacrifice, and captivity introduces fragility and confinement. The caged-bird reading is reinforced by lore that frames consciousness and memory through enclosure. The character therefore fuses professional, literary, and affective traces into a single tragic healer.

Warfarin uses prototype contradiction. The vampire trace is immediately legible through blood imagery, pale features, red eyes, and nocturnal associations. Yet her name is also a medical term, and her role as a medic shifts the vampire figure from predation to care. The design creates tension between blood as danger and blood as medical knowledge. This is one of the strongest examples of prototype resonance because the character depends on the clash between two cultural domains: gothic horror and clinical treatment.

Czerny relies on historical-professional recognition. The piano, musical background, and name evoke the world of classical music and the pedagogical memory associated with Czerny. His character identity emerges through disciplined craft, cultural prestige, and the circulation of music across community life. In applied-linguistic terms, the character is built

through a professional register as much as through fantasy imagery: the music world supplies the semiotic resources for identity.

Hung activates a guardian prototype. The image links his design to Komainu and to the wider movement of stone-lion guardian imagery across East Asian religious and architectural cultures. His shield, animal features, and kind but protective disposition make guardianship central. The prototype trace here is not a single historical figure but a cultural image type. This matters because games often draw not only on famous individuals but on social images that have circulated across everyday visual culture.

Skadi is organized through oceanic and siren-like resonance. The image connects the character to the mythological siren, especially the relation between song, sea, danger, and melancholy. Blue coloration, sea-monster background imagery, and lore about haunting song create a figure whose narrative identity is carried by sound and ocean before it is explained by plot. Skadi shows that prototype traces can be sonic in implication even when visually represented.

Pallas offers a classical example. Her name, olive leaves, ritual symbols, references to Minos, sacred fire, and heroic diction all point toward Greek myth and Athena-like associations. The character is not simply Greek-themed. Her multimodal design translates priesthood, blessing, exile, heroic memory, and collective morale into a combat-support identity. This is character-narrative coupling at work: the visual design prepares the player to understand the character's role in story, combat, and cultural world-building.

Across these cases, a recurring pattern appears. The image proposes a prototype; the text confirms, redirects, or complicates it; the game role makes it playable. The resulting character is a multimodal construction.

3.3.6 From Character Image to Worldview

The cumulative effect of these character designs is worldview construction. *Arknights*

presents a world where mythological beings, medical institutions, military technologies, folklore traces, and historical professions can coexist. The world is assembled through repeated character encounters. Each operator adds a semiotic fragment: vampire medicine, Greek ritual, gothic theatre, East Asian guardianship, Norse fire, oceanic song, revolutionary protection.

This is why multimodal discourse analysis is useful for game narrative. If analysis focuses only on plot, it may miss how the game world is prepared by character images. If analysis focuses only on mechanics, it may miss how role and action are narratively justified. If analysis focuses only on localization or translation, it may miss how visual-cultural prototypes cross linguistic boundaries. A multimodal approach can show how the game constructs narrative coherence through distributed signs.

4. Discussion

4.1 Toward a Gameplay Visual Grammar

The findings suggest that visual grammar can be extended toward what might be called a gameplay visual grammar. In a traditional image, representational, interactive, and compositional meanings organize how viewers understand the image. In a game, these meanings also organize possible action. The character's gaze invites address; the body and equipment invite evaluation; the center-margin composition supports selection and recognition; symbolic attributes prepare combat role and narrative expectation.

This means that Kress and van Leeuwen's framework should be adaptably applied to games with modification. Games add procedural and interactive dimensions that visual grammar alone cannot fully capture. A static portrait becomes meaningful in relation to recruitment systems, battle deployment, voice playback, upgrade interfaces, and player attachment. Nevertheless, visual grammar remains a powerful entry point because many game actions begin with visual reading. The player must first recognize what kind of figure is being offered.

The *Arknights* corpus shows that symbolic processes are particularly important in character-collection games. A symbolic portrait can make a new operator understandable at a glance. The portrait is not a full narrative, but it is a narrative premise. It says: this figure belongs to a world of fire, blood, medicine, opera, guardianship, music, or classical ritual. The player then reads subsequent texts in light of that premise.

4.2 Prototype Resonance as Multimodal Intertextuality

The case readings also suggest that game intertextuality is often multimodal rather than purely verbal. Fairclough's (1992) account of intertextuality describes how texts are shaped by prior texts. In games, prior texts may return as color palettes, poses, props, musical implications, species designs, or interface metaphors. A player may encounter a trace of Dracula without the novel being named; a trace of Athena without a direct mythological explanation; a trace of The Phantom of the Opera through mask, theater, and French allusion.

Forceville and Urios-Aparisi's (2009) work on multimodal metaphor helps clarify this process. A character such as Warfarin is not merely a vampire plus a doctor. The design maps one domain onto another: blood becomes both gothic substance and medical object; danger becomes care; monstrosity becomes professional competence. Similarly, Nightingale maps healing onto captivity and sacrifice. They shape how the player interprets agency, vulnerability, and moral position.

Prototype resonance is therefore a useful addition to multimodal discourse analysis. It names the player's uptake of culturally sedimented signs. It also helps explain why some game characters feel immediately "coherent" even before the player has read extensive lore. Coherence comes from the successful alignment of prototype, visual grammar, textual framing, and game function.

4.3 Applied Linguistic Implications: Game Literacy and Cross-Cultural Reading

The study has implications for applied linguistics in at least three areas. First, it expands

the object of multimodal discourse analysis. Digital games should be treated not only as sites of language learning or online communication, but as complex discourse artifacts whose visual, textual, and procedural signs construct meaning. This aligns with multiliteracies theory, which treats literacy as design across modes (New London Group, 1996).

Second, the study highlights the importance of critical game literacy. Players learn to interpret characters through cultural prototypes, but those prototypes are not neutral. They carry histories of gender, ethnicity, profession, mythology, morality, and genre. A critical reader of games should be able to ask how a character design transforms its sources, what it makes visible, what it simplifies, and what cultural knowledge it assumes.

Third, the findings matter for localization and intercultural communication. A prototype trace may be highly legible in one cultural context and weaker in another. Hung's guardian imagery may resonate differently for players familiar with East Asian temple and architectural guardians than for players who encounter the image only as a fantasy animal-person design. Pallas's classical Greek imagery may be read as myth, as heroic antiquity, or simply as an exoticized visual style. Localization is therefore a mediation of multimodal recognizability.

4.4 Methodological Implications for Multimodal Game Analysis

The study also points toward a methodological lesson. Game images should not be isolated from their interface and textual environment. A character portrait changes meaning when placed in a recruitment menu, a battle formation screen, an operator file, or an event story. Similarly, an interface button can become narratively meaningful when it is embedded in a world-internal role. Multimodal game analysis must therefore move between micro-level visual details and macro-level discourse functions.

The three-stage procedure used here - visual grammar coding, prototype trace identification, and character-narrative coupling - is one possible method. It offers a way to preserve the precision of visual analysis while still explaining why visual features matter for

story and play. Future studies could add player reception data, eye-tracking, think-aloud protocols, or cross-cultural comparison to test how different players actually take up prototype traces. The present article offers a possible framework; reception research would deepen it.

5. Conclusion

5.1 Summary of Contributions

This paper has focused on videogame character image construction and game narrative construction. It has made three main contributions.

First, it introduced the concept of multimodal prototype analysis. These analysis describe the cultural residues embedded in game character design through image, name, prop, color, layout, voice, and textual framing. They explain how players can infer character identity before full narrative exposition.

Second, it proposed prototype resonance as the process through which players connect those traces to prior cultural knowledge. Prototype resonance is is the interpretive movement by which an older cultural repertoire becomes useful for reading a new gameplay character.

Third, it argued that the interface and character images of *Arknights* produce character-narrative coupling. The interface turns routine operations into diegetic actions performed by the Doctor, while character portraits convert myth, history, folklore, genre, and profession into playable identities. Visual grammar is not only an analytic vocabulary for images, but also a way to understand how images invite action and narrative participation.

5.2 Limitations and Future Research

The study has limitations. The analysis does not include player interviews, survey data, or reception evidence, so claims about player interpretation remain theoretically grounded in the affordances of the game materials rather than directly observed player responses.

Future research could compare *Arknights* with other character-collection games, role-playing games, or multiplayer online games. It could also examine how players from different

linguistic and cultural backgrounds interpret the same prototype traces. Another useful direction would be to analyze voice, music, animation, and combat mechanics alongside static character portraits. Such work would move from visual grammar toward a fuller multimodal discourse framework of gameplay experiences.

References

- Aarseth, E. J. (1997). *Cybertext: Perspectives on ergodic literature*. Johns Hopkins University Press.
- Consalvo, M., & Dutton, N. (2006). Game analysis: Developing a methodological toolkit for the qualitative study of games. *Game Studies*, 6(1). https://gamestudies.org/0601/articles/consalvo_dutton
- Ensslin, A. (2011). *The language of gaming*. Palgrave Macmillan. <https://doi.org/10.1007/978-0-230-35708-2>
- Fairclough, N. (1992). *Discourse and social change*. Polity Press.
- Forceville, C., & Urios-Aparisi, E. (Eds.). (2009). *Multimodal metaphor*. Mouton de Gruyter. <https://doi.org/10.1515/9783110215366>
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan. <https://doi.org/10.1145/950566.950595>
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge. <https://doi.org/10.4324/9780203431269>
- Jenkins, H. (2004). Game design as narrative architecture. In N. Wardrip-Fruin & P. Harrigan (Eds.), *First person: New media as story, performance, and game* (pp. 118–130). MIT Press.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge. <https://doi.org/10.4324/9781315638027>
- Juul, J. (2005). *Half-real: Video games between real rules and fictional worlds*. MIT Press.
- Kress, G. (2003). *Literacy in the new media age*. Routledge. <https://doi.org/10.4324/9780203164753>
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge. <https://doi.org/10.4324/9780203970034>

- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge. <https://doi.org/10.1080/15551393.2021.1992242>
- O'Halloran, K. L. (Ed.). (2004). *Multimodal discourse analysis: Systemic functional perspectives*. Continuum.
- Ryan, M.-L. (2006). *Avatars of story*. University of Minnesota Press. <https://doi.org/10.5749/9780816697861>
- Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. MIT Press.
- The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60–93. <https://doi.org/10.17763/haer.66.1.17370n67v22j160u>
- Thorne, S. L., Black, R. W., & Sykes, J. M. (2009). Second language use, socialization, and learning in Internet interest communities and online gaming. *The Modern Language Journal*, 93(s1), 802–821. <https://doi.org/10.1111/j.1540-4781.2009.00974.x>
- Yostar Limited. (2019). *Arknights* [Mobile game]. <https://www.Arknights.global/>

Appendix A

Table A1

Codebook for the Analysis of Character Image and Narrative Construction

Code node	Conceptual definition	Typical evidence	Boundary notes
Symbolic representation	Character is presented as bearer of attributes rather than as participant in a depicted action	Staff, sword, shield, mask, animal feature, ritual object	Distinguished from narrative process, where vectors of action dominate the image
Demand address	Character gaze creates a relation with the viewer/player	Direct eye contact, frontal position, assistant-like address	Distinguished from offer images, where the figure is presented for contemplation
Prototype trace	Visual, verbal, or cultural cue evokes a recognizable prior repertoire	Vampire imagery, Greek ritual signs, opera mask, guardian-animal motif	Requires relevance to character interpretation, not merely superficial resemblance
Prototype resonance	Prototype trace shapes interpretation of role, morality, mood, or narrative expectation	Warfarin read through vampire medicine; Pallas through Athena-like ritual heroism	Distinguished from source identification alone; the trace must affect character reading
Character-narrative coupling	Image, lore, and gameplay role reinforce a coherent character identity	Healer image plus medical role plus sacrificial lore	Includes productive tension when contradiction is narratively meaningful
Operational narration	Interface action is framed as a world-internal action	Login as reconnection, menu as command center, battle selection as tactical decision	Distinguished from purely technical interface functions