



A Multimodal Analysis of the Makuku Diapers Commercial Video Featuring Nikita Willy

Frina Diniarta Nur

frinadiniarta.nur@gmail.com

English Department, Institut Bisnis Nusantara, Jakarta

Rini Adriati¹

rini_adriati@yahoo.com

English Department, Institut Bisnis Nusantara, Jakarta

Lavenia Oktaviani

laveniaoktaviani01@gmail.com

English Department, Institut Bisnis Nusantara, Jakarta

Nur, F.D., Adriati, R. and Oktaviani, L. (2026). A Multimodal Analysis of the Makuku Diapers Commercial Video Featuring Nikita Willy. *Journal of English Language and literature*, 11(1), 153-168. doi: 10.37110/jell.v11i1.343

Received: 10-01-2026

Accepted: 20-02-2026

Published: 02-03-2026

Abstract: Advertising increasingly relies on multimodal resources to communicate emotional and persuasive messages effectively. However, studies examining how multimodal elements function collectively in Indonesian television commercials, particularly those featuring local celebrities, remain limited. This study aims to analyze the use and effectiveness of multimodal elements in the Makuku diapers commercial video featuring Nikita Willy. A qualitative multimodal discourse analysis was employed to examine five modes: linguistic, visual, gestural, spatial, and audio. The data were collected from the official commercial video and analyzed using a social semiotic framework. The effectiveness of each mode was evaluated using a researcher-defined three-level coding scheme: effective, somewhat effective, and ineffective. The findings reveal that all five multimodal modes are systematically integrated to construct emotional appeal and reinforce the brand message. The visual, gestural, and spatial modes are highly effective in creating intimacy, warmth, and family connection, while the linguistic and audio modes support memorability and emotional engagement, although some aspects, particularly sound effects and secondary textual elements, show limited prominence. Overall, the coordinated use of multimodal resources strengthens the commercial's persuasive impact by aligning emotional representation with brand identity. These findings highlight the importance of strategic multimodal integration in enhancing advertising effectiveness and provide practical insights for advertising creators and researchers examining multimodal communication in commercial contexts.

Keywords: *multimodal discourse analysis, advertising, social semiotics, commercial video, Indonesian advertising*

INTRODUCTION

Communication plays a central role in human interaction, enabling individuals and organizations to convey ideas, emotions, and

intentions across various contexts. In contemporary digital environments, communication increasingly occurs through multiple semiotic resources, including language,

¹ Corresponding author

images, sound, gesture, and spatial arrangement. These resources function as meaning-making systems that allow communicators to construct messages and influence audience interpretation. As media technologies continue to evolve, communication is no longer limited to written or spoken language but involves the integration of multiple modes working simultaneously to create coherent and persuasive messages. This shift reflects broader transformations in digital communication, where multimodal design is strategically employed to enhance audience engagement, emotional resonance, and persuasive effectiveness (Kress, 2010; van Leeuwen, 2005).

Semiotics, defined as the study of signs and meaning-making processes, provides a theoretical foundation for understanding how meaning is constructed and interpreted. According to Ferdinand de Saussure (1916/1983), meaning is produced through the relationship between the signifier and the signified, forming the basis of communication. This perspective highlights that meaning is not inherent in objects themselves but is created through socially and culturally shared systems of signs. In media communication, semiotic analysis enables researchers to examine how various sign systems—including linguistic, visual, and auditory elements—work together to convey messages and influence audience perception (Chandler, 2007). In digital advertising contexts, semiotics is particularly relevant because advertisements rely heavily on symbolic representation, visual identity, and emotionally resonant imagery to shape audience interpretation and brand perception.

Building on semiotic theory, multimodality emphasizes that contemporary communication involves the integration of multiple semiotic modes rather than relying on a single mode. According to Kress and van Leeuwen (2001, 2006), modes such as language, image, gesture, spatial arrangement, and sound function as semiotic resources, each contributing uniquely to meaning-making. Multimodal analysis examines how these modes interact to produce coherent and meaningful messages. This approach has become increasingly important in digital advertising, where meaning is constructed through the coordinated interaction of visual design, spoken language, background music, bodily expression, and spatial composition. These multimodal configurations not only convey information but also shape emotional engagement, audience

attention, and persuasive impact (Kress, 2010; van Leeuwen, 1999).

Recent developments in digital media have further expanded the relevance of multimodal discourse analysis, particularly in advertising distributed through online platforms such as YouTube, social media, and streaming services. Digital advertisements integrate audiovisual elements, branding visuals, and narrative storytelling techniques to create immersive and emotionally engaging experiences. Scholars have noted that multimodal design in digital advertising plays a crucial role in shaping audience perception, strengthening brand identity, and enhancing persuasive communication (Jewitt, 2014; Bateman, Wildfeuer, & Hiippala, 2017). These developments highlight the importance of analyzing not only the presence of multimodal elements but also how effectively these elements function together to support communicative and persuasive goals in contemporary advertising contexts.

While semiotics can examine meaning within individual modes, real-world media texts, particularly video commercials, rely on the interaction of multiple modes simultaneously. Multimodality refers to the combined use of linguistic, visual, gestural, spatial, and audio modes to construct meaning. Each mode contributes distinct communicative affordances, and their integration enhances the overall clarity, emotional appeal, and persuasive effectiveness of the message (Kress & van Leeuwen, 2006; van Leeuwen, 1999). In video advertising, spoken language conveys explicit information, visual imagery attracts attention and builds brand identity, gestures express emotions and interpersonal meaning, spatial arrangement guides viewer focus, and audio elements reinforce mood and emotional engagement. The interaction of these modes creates a cohesive and persuasive communicative experience.

Several previous studies have applied multimodal discourse analysis to media texts, demonstrating the value of this approach in understanding meaning construction. Muliani and Nur Ali (2022), for example, analyzed the music video “Lathi” using multimodal discourse analysis to examine how visual, gestural, and audio modes contribute to interpersonal meaning. Their findings showed that multimodal integration effectively conveyed emotional

transitions and cultural identity. However, while the study provided valuable insights into meaning construction, its focus on music videos rather than commercial advertising limits its applicability to persuasive and branding contexts. In addition, the study primarily examined meaning representation without systematically evaluating how individual multimodal elements contribute to communicative effectiveness.

Similarly, Muthaharoh (2020) examined verbal and visual elements in the OPPO Reno3 Pro television advertisement using multimodal discourse analysis. The findings demonstrated that verbal and visual modes worked together to create a cohesive promotional message and attract viewer attention. However, the study focused primarily on two modes—verbal and visual—without examining other important modes such as gesture, spatial arrangement, and audio. As a result, the analysis did not fully capture the complexity of multimodal interaction in video advertising, where meaning is constructed through the integration of multiple semiotic resources.

Masrukin (2023) also analyzed multimodal meaning in a Pantene advertisement using Systemic Functional Linguistics and multimodal discourse theory. The study showed that verbal and visual elements effectively contributed to meaning construction and audience engagement. However, the research focused mainly on representational meaning and did not systematically assess the effectiveness of individual multimodal aspects in achieving communicative and persuasive objectives. This highlights the need for further research that not only identifies multimodal elements but also evaluates their effectiveness in supporting advertising goals.

These previous studies demonstrate that multimodal discourse analysis provides valuable insights into meaning construction in media texts. However, several aspects remain insufficiently explored, particularly in the context of contemporary Indonesian digital advertising. First, while previous research has examined multimodal meaning in advertisements, relatively few studies have analyzed commercial video advertisements using a comprehensive framework that simultaneously examines linguistic, visual, gestural, spatial, and audio modes. Second, prior studies have tended to focus on meaning representation rather than

systematically evaluating how effectively each multimodal aspect contributes to communicative clarity, emotional engagement, and persuasive impact. Third, with the increasing prominence of digital video advertising, there is a growing need to examine how multimodal resources function strategically to enhance audience engagement and strengthen brand identity in digital media environments.

To address these gaps, the present study analyzes the Makuku diapers commercial video featuring Nikita Willy using a multimodal analytical framework based on Kress & van Leeuwen (2001, 2006), van Leeuwen (1999, 2006, 2011), van Leeuwen & Jewitt (2001), and McNeill (1992). This study examines how five multimodal modes—linguistic, visual, gestural, spatial, and audio—interact to construct meaning and support the advertisement's communicative and persuasive objectives. In addition, this study applies a systematic analytical classification to evaluate the effectiveness of each multimodal aspect in enhancing message clarity, emotional appeal, audience engagement, and brand identity.

The novelty of this study lies not only in its focus on an Indonesian digital commercial but also in its systematic evaluation of multimodal effectiveness across multiple semiotic modes. By examining how linguistic, visual, gestural, spatial, and audio modes function individually and interactively, this study provides a more comprehensive understanding of how multimodal resources contribute to persuasive communication in digital advertising. Furthermore, the use of an effectiveness-based analytical framework offers a structured approach to assessing the communicative contribution of each mode, extending previous multimodal research that has primarily focused on meaning representation rather than communicative effectiveness.

This research contributes to multimodal discourse analysis by providing empirical insights into how multimodal elements function strategically in contemporary digital advertising. It also contributes to advertising and communication studies by demonstrating how multimodal resources enhance communicative clarity, emotional engagement, and persuasive effectiveness. More broadly, this study supports the growing recognition of multimodality as a key analytical framework for understanding meaning-making in digital media communication.

METHOD

This study employed a descriptive qualitative method, which is appropriate for semiotic and multimodal analysis because it allows an in-depth interpretation of signs, meanings, and their functions within a specific communicative context. This method enabled the writers to analyze how linguistic, visual, gestural, spatial, and audio modes contribute to meaning-making and persuasive communication in an advertisement video

The primary data source was a commercial video published on the official YouTube channel MAKUKU Indonesia Official on the YouTube platform. The video analyzed was a Makuku diapers commercial featuring Nikita Willy, accessible at:

https://www.youtube.com/watch?v=BKTN7P1jCWs&ab_channel=MAKUKUIndonesiaOfficial

The unit of analysis consisted of multimodal elements appearing throughout the video, including spoken text, written text, visual imagery, gestures, spatial arrangements, and audio components. These elements were examined to answer six research questions related to semiotic signs and the communicative functions of each multimodal mode

To analyze the data, the writers applied multimodal social semiotic theories. The identification of semiotic signs was based on the framework of Kress and van Leeuwen (2001). The linguistic mode was analyzed using the frameworks of Kress and van Leeuwen (2001) and van Leeuwen (2006), focusing on text structure, typography, and layout. The visual mode was analyzed using the theories of Kress and van Leeuwen (2006), van Leeuwen (2011), and van Leeuwen and Jewitt (2001), examining colour, style, and perspective. The gestural mode was analyzed using the theory proposed by McNeill (1992), focusing on facial expressions, body language, and interpersonal interaction. The spatial mode was analyzed using Kress and van Leeuwen (2006) and van Leeuwen (2006), focusing on placement, directional flow, and proximity. The audio mode was analyzed using van Leeuwen's (1999) framework, examining music, sound effects, silence, and tone.

In addition to identifying and interpreting the multimodal elements, the writers classified the

effectiveness of each mode and its aspects in supporting the advertisement's communicative and persuasive goals. The effectiveness of each multimodal aspect was evaluated using a researcher-assigned three-level analytical classification consisting of Effective, Somewhat Effective, and Ineffective. These categories functioned as a coding framework within qualitative content analysis, enabling a systematic and consistent evaluation of how linguistic, visual, gestural, spatial, and audio elements contributed to message clarity, audience engagement, emotional appeal, and the reinforcement of brand identity.

To ensure analytical consistency and transparency, explicit criteria were established to evaluate the effectiveness of each multimodal aspect in supporting the advertisement's communicative and persuasive goals.

The effectiveness of each multimodal aspect was evaluated using a researcher-assigned three-level analytical classification based on explicit communicative and semiotic criteria, as follows:

Effective

An aspect was classified as Effective when it clearly and strongly supported the advertisement's communicative and persuasive goals. This included enhancing message clarity, reinforcing brand identity, increasing audience engagement, and contributing meaningfully to the emotional and symbolic impact of the advertisement.

Somewhat Effective

An aspect was classified as Somewhat Effective when it partially supported the communicative and persuasive goals but had limitations in visibility, clarity, salience, or emotional impact. While the aspect contributed to meaning-making, its communicative strength was not fully optimal.

Ineffective

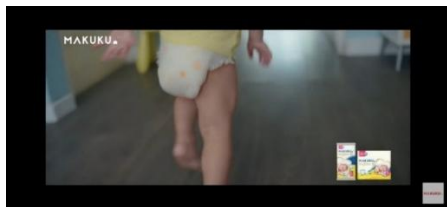
An aspect was classified as Ineffective when it did not support, or only minimally supported, the communicative and persuasive goals of the advertisement. Such aspects did not significantly contribute to message clarity, brand reinforcement, audience engagement, or emotional expression.

The classification process involved careful observation, repeated viewing of the video,

identification of relevant multimodal elements, interpretation based on established multimodal frameworks, and systematic categorization using the defined criteria. This procedure ensured analytical transparency and consistency in evaluating how multimodal resources functioned to convey meaning and persuasion in the commercial video.

FINDINGS AND DISCUSSION

To facilitate the analysis of semiotic signs in this commercial video, the writers divided the video into several segments, referred to as data points. In total, there are 15 data points. These data were then analyzed using the five modes of multimodality: linguistic, visual, gestural, spatial, and audio. Each mode was examined by observing its supporting aspects to determine whether each aspect was used effectively. The linguistic mode was analyzed through text structure, typography, and layout. The visual mode was examined based on color, style, and perspective. The gestural mode focused on facial expressions, body language, and interactions. The spatial mode considered the placement, organization, and proximity of people and objects. The audio mode was analyzed based on music, sound effects, silence, and tone. The following section presents a sample analysis of three representative data points.



Picture 1. Baby Issa is running in the hallway to his mother

Context

This screenshot was captured at 0.5 seconds of a Makuku diapers commercial video featuring Nikita Willy and her son, Baby Issa. In this scene, Baby Issa is shown running along a hallway toward his mother, who remains off-screen, while wearing a yellow outfit and a Makuku Dry Care diaper. The camera captures the baby's movement from behind, visually emphasizing the diaper and his unrestricted mobility. Accompanying this visual scene is a voiceover in the form of a jingle that states: "Makuku, mamaku, aku sayang kamu, Makuku ku ku ku ku sayangku. Makuku, papaku, aku sayang kamu, Makuku ku ku ku ku sayangku." This combination of visual and auditory elements

establishes an emotionally warm and affectionate atmosphere that reflects the familial theme of the advertisement

Analysis

This data was analyzed using a multimodal social semiotic framework based on the theories of Kress & van Leeuwen (2001, 2006), van Leeuwen (1999, 2006, 2011), van Leeuwen & Jewitt (2001), and McNeill (1992). This framework examines how meaning is constructed through multiple semiotic modes. The analysis focuses on five modes: linguistic, visual, gestural, spatial, and audio. Each mode was examined to identify its semiotic resources and evaluate its effectiveness in supporting the advertisement's communicative and persuasive goals.

1. Linguistic Mode

The effectiveness of the linguistic mode was analyzed using the frameworks of Kress and van Leeuwen (2001) and van Leeuwen (2006), focusing on text structure, typography, and layout. These semiotic resources were examined to determine their contribution to message clarity, emotional engagement, brand reinforcement, and the advertisement's persuasive objectives.

In terms of text structure, the scene is classified as somewhat effective. The spoken language in the jingle, such as "Makuku, mamaku, aku sayang kamu" and "sayangku," employs repetition and affectionate expressions that enhance memorability and emotional resonance. According to Kress and van Leeuwen (2001), linguistic structure functions as a semiotic resource for conveying interpersonal meaning and shaping audience relationships with the message. The repetition of affectionate expressions establishes a strong emotional association between the brand and familial love, reinforcing the advertisement's nurturing and caring theme. This linguistic strategy contributes positively to emotional engagement and supports the persuasive objective of positioning the brand within a warm family context.

However, the written linguistic elements contribute less effectively to communicative clarity and brand reinforcement. While the Makuku logo in the upper-left corner is clearly visible and legible (see Picture 1), the smaller logo and product text in the bottom-right corner have lower visual salience due to their reduced size and peripheral placement. According to Kress and van Leeuwen (2001), the effectiveness of linguistic signs depends not only on their presence but also on their visibility and

integration within the overall semiotic composition. The limited prominence of the smaller written elements reduces their ability to reinforce brand identity and clearly communicate product information. Although the smaller logo also functions as an interactive link to the brand's official YouTube channel, its low salience weakens its communicative impact within the scene. Therefore, while the spoken linguistic elements effectively support emotional engagement, the written linguistic elements only partially support message clarity and brand reinforcement, resulting in an overall classification of somewhat effective.

In terms of typography, the scene is classified as somewhat effective. The Makuku logo uses a clean, white sans-serif typeface that contrasts clearly with the background, ensuring legibility and supporting immediate brand recognition (see Picture 1). Kress and van Leeuwen (2001) explain that typography functions as a semiotic resource by conveying meaning through visual clarity, stylistic form, and associative value. The simple and modern typeface reinforces a trustworthy and professional brand identity, which supports the advertisement's persuasive objective and enhances brand credibility. This clear typographic presentation contributes positively to communicative clarity and brand recognition.

However, the smaller logo and accompanying product text in the bottom-right corner have reduced size and visual prominence, limiting their readability and communicative effectiveness. Their lower salience reduces their contribution to reinforcing brand identity and communicating product-related information clearly. As a result, although the primary logo is typographically effective in supporting brand recognition, the overall typographic contribution within the scene only partially fulfills the criteria of message clarity and brand reinforcement. Therefore, the typographic aspect is classified as somewhat effective.

In terms of layout, the scene is classified as somewhat effective. The primary Makuku logo is positioned in the upper-left corner, which aligns with the conventional reading path and increases its visual salience (see Picture 1). According to Kress and van Leeuwen (2001), layout organizes the spatial arrangement of semiotic resources and plays a crucial role in directing viewer attention and structuring meaning. The prominent placement of the logo enhances its visibility and supports brand recognition, contributing

positively to communicative clarity and persuasive effectiveness.

However, the smaller logo and product images in the bottom-right corner receive less visual attention due to their reduced size and peripheral location. This uneven distribution of visual emphasis limits their ability to function effectively as linguistic signs that reinforce brand identity and communicate product information. Kress and van Leeuwen (2001) emphasize that effective layout requires strategic placement to ensure optimal visibility and meaning-making. In this scene, the reduced prominence of secondary linguistic elements weakens their contribution to the advertisement's communicative and persuasive goals. Consequently, while the primary logo placement supports brand recognition, the overall layout only partially fulfills the criteria of communicative clarity and brand reinforcement, resulting in a classification of somewhat effective.

2. Visual Mode

The effectiveness of the visual mode was analyzed using the theories of Kress and van Leeuwen (2006), van Leeuwen (2011), and van Leeuwen and Jewitt (2001), focusing on colour, style, and perspective. These semiotic resources were examined to determine their contribution to communicative clarity, emotional engagement, brand reinforcement, and the advertisement's persuasive goals.

Regarding the first aspect of the visual mode, the color choice in the scene is classified as effective. The scene presents a warm and inviting atmosphere through the use of soft lighting and harmonious color combinations, which enhance the sense of comfort within the home environment. According to Kress and van Leeuwen (2006), color functions as a semiotic resource that conveys emotional meaning and shapes audience perception. The white color of the Makuku logo and the diaper worn by Baby Issa signifies purity, cleanliness, and safety, which are desirable qualities associated with baby care products. This visual representation reinforces the product's identity as hygienic and reliable, supporting brand reinforcement and communicative clarity.

In addition, Baby Issa's yellow outfit conveys cheerfulness, warmth, and positive emotional energy. Van Leeuwen (2011) explains that color can evoke emotional associations that contribute to interpersonal meaning and viewer engagement. The yellow color enhances the emotional appeal

of the scene and strengthens the association between the product and positive childhood experiences. Furthermore, the brown floor tiles, which resemble natural wood, evoke warmth, stability, and a nurturing home environment. This combination of colors creates a visually cohesive and emotionally comforting atmosphere that aligns with the advertisement's themes of family love, safety, and happiness. Therefore, the color choices effectively enhance emotional engagement, reinforce brand identity, and support the advertisement's persuasive objective.

Regarding the second aspect of the visual mode, the scene's style is classified as effective. The room and hallway are characterized by clean, straight lines and a simple, uncluttered design. According to van Leeuwen and Jewitt (2001), visual style functions as a semiotic resource that communicates social meaning, including values such as order, modernity, and practicality. The tidy and organized environment conveys a sense of cleanliness, control, and safety, which are important qualities in the context of baby care products.

The use of minimal decoration ensures that visual attention remains focused on the baby and the product, rather than being distracted by unnecessary background elements. This stylistic choice enhances communicative clarity by directing viewer attention toward the central message of the advertisement. In addition, the warm-toned flooring contributes to a comforting and nurturing environment, reinforcing the emotional theme of family care. This visual style supports persuasive effectiveness by creating a credible and relatable domestic setting that aligns with the product's intended use. Therefore, the visual style effectively enhances message clarity, emotional engagement, and brand credibility.

Regarding the third aspect of the visual mode, the scene's perspective is classified as effective. The visual composition places Baby Issa at the central focal point of the scene, with the camera angle capturing his movement clearly and naturally. According to Kress and van Leeuwen (2006), perspective functions as a semiotic resource that structures viewer attention and establishes relationships between represented participants and viewers. The central positioning of Baby Issa increases his visual salience, ensuring that viewers immediately recognize him as the primary subject of the scene.

The visual emphasis on Baby Issa's white diaper and cheerful yellow outfit reinforces the product's visibility and symbolic meaning. This

strategic visual focus strengthens brand reinforcement by directly associating the product with positive emotional experiences such as safety, comfort, and happiness. In addition, the perspective highlights the baby's unrestricted movement, which implicitly communicates the product's functional benefit in supporting mobility and comfort. This visual representation enhances communicative clarity by demonstrating the product's practical value while simultaneously increasing emotional engagement through the portrayal of a joyful and active child. Therefore, the perspective effectively supports the advertisement's communicative and persuasive goals.

3. Gestural Mode

The effectiveness of the gestural mode was analyzed using McNeill's (1992) theory, focusing on facial expressions, body language, and interpersonal interaction. Gestures function as meaningful bodily expressions that convey emotional states, intentions, and interpersonal relationships. These semiotic resources were examined to determine their contribution to communicative clarity, emotional engagement, brand reinforcement, and the advertisement's persuasive goals.

Regarding the first aspect of the gestural mode, the facial expressions in the scene are classified as ineffective. The camera captures Baby Issa from behind, showing only his back and diaper, while his face remains outside the viewer's visual field. According to McNeill (1992), facial expressions function as important semiotic resources that convey emotional meaning and support interpersonal communication. Facial expressions help viewers interpret emotional states and establish emotional connections with represented participants.

In this scene, the absence of visible facial expressions limits the viewer's ability to directly observe the baby's emotional response. As a result, this aspect does not significantly contribute to emotional engagement or enhance communicative clarity regarding the baby's feelings. Although other semiotic modes, such as body language and audio, help convey emotional meaning, the facial expression aspect itself does not directly support the advertisement's communicative and persuasive goals. Therefore, this aspect is classified as ineffective.

Regarding the second aspect of the gestural mode, the body language in the scene is classified as effective. Baby Issa's energetic and natural

running movement along the hallway visually communicates freedom of movement and physical comfort. According to McNeill (1992), body gestures function as semiotic resources that express action, intention, and emotional states, contributing to meaning-making beyond spoken language.

The baby's smooth and unrestricted movement implicitly demonstrates the functional benefit of the Makuku Dry Care diaper, particularly its comfort and flexibility. This visual representation enhances communicative clarity by showing, rather than verbally explaining, the product's effectiveness in supporting natural movement. In addition, the baby's playful and confident body language contributes to emotional engagement by conveying a sense of happiness, safety, and well-being. This strengthens the persuasive impact of the advertisement by associating the product with positive physical and emotional experiences. Therefore, the body language effectively supports message clarity, emotional engagement, and persuasive effectiveness.

Regarding the third aspect of the gestural mode, the interaction between people in the scene is classified as effective. Although Nikita Willy is not directly visible on screen, Baby Issa's forward movement toward an off-screen participant indicates an ongoing interpersonal interaction. According to McNeill (1992), gestures function as semiotic resources that help construct interpersonal relationships and communicate social meaning. The baby's movement toward his mother suggests intentional interaction, reflecting emotional connection and relational closeness.

This implied interaction contributes significantly to emotional engagement by portraying a nurturing and affectionate parent-child relationship. The viewer can infer the presence of a caring and supportive environment, which aligns with the advertisement's thematic focus on family love, protection, and trust. This emotional representation reinforces brand identity by associating the product with parental care and child well-being. In addition, the interaction supports persuasive effectiveness by presenting the product within a meaningful and relatable family context. Therefore, the gestural interaction effectively contributes to emotional engagement, brand reinforcement, and the advertisement's overall communicative goals.

4. Spatial Mode

The effectiveness of the spatial mode was analyzed using Kress and van Leeuwen (2006) and van Leeuwen (2006), focusing on placement, directional flow, and proximity. Spatial organization functions as a semiotic resource that structures meaning by guiding viewer attention, establishing relationships between elements, and supporting narrative coherence. These semiotic resources were examined to determine their contribution to communicative clarity, emotional engagement, brand reinforcement, and the advertisement's persuasive goals.

Regarding the first aspect of the spatial mode, the placement of elements in the scene is classified as effective. Baby Issa is positioned centrally within the frame while running along the hallway, making him the most visually salient participant. According to Kress and van Leeuwen (2006), spatial placement influences visual salience and determines which elements viewers perceive as most important. The central positioning ensures that viewers immediately focus on the baby, which enhances communicative clarity by clearly presenting the primary subject of the advertisement.

In addition, the Makuku logo is positioned in the upper-left corner, a location associated with high visual visibility due to conventional reading patterns. This placement supports brand reinforcement by ensuring that the brand identity remains consistently visible without interfering with the viewer's focus on the main action. The smaller logo and product images placed in the bottom-right corner provide additional product visibility while maintaining visual balance within the composition. Van Leeuwen (2006) explains that spatial arrangement helps organize semiotic resources in ways that support meaning-making and viewer interpretation. The balanced placement of the baby, logos, and product images ensures that viewers can simultaneously recognize the product and emotionally engage with the narrative. Therefore, the placement of elements effectively enhances message clarity, brand reinforcement, and persuasive effectiveness.

Regarding the second aspect of the spatial mode, the directional flow in the scene is classified as effective. The sequence of visual movement begins with the appearance of Baby Issa, followed by the camera capturing the hallway and then focusing on his running movement. According to Kress and van Leeuwen (2006), directional flow functions as a semiotic

resource that guides viewer attention and structures the narrative progression of visual communication. This movement creates a clear visual pathway that allows viewers to follow the unfolding action in a logical and coherent manner.

The directional flow emphasizes the baby's movement toward an implied destination, which contributes to narrative coherence and emotional engagement. This progression strengthens communicative clarity by helping viewers understand the sequence of events and the relational context between the baby and his environment. In addition, the movement enhances persuasive effectiveness by visually demonstrating the baby's comfort and mobility while wearing the product. Therefore, the directional flow effectively supports narrative clarity, emotional engagement, and the advertisement's communicative goals.

Regarding the third aspect of the spatial mode, the proximity between people and objects in the scene is classified as effective. Baby Issa is shown moving within a clean and familiar home environment, with spatial relationships that emphasize his connection to the surrounding domestic space. According to van Leeuwen (2006), proximity functions as a semiotic resource that conveys interpersonal relationships and social meaning by visually representing closeness or distance between participants and their environment.

The close spatial relationship between the baby and the home environment reinforces the sense of safety, comfort, and belonging. This spatial arrangement enhances emotional engagement by presenting a nurturing and supportive domestic setting. In addition, the hallway visually guides the viewer's attention toward the implied presence of the mother, reinforcing the relational and emotional context of family care. This spatial proximity also supports brand reinforcement by associating the product with a safe and caring home environment. Therefore, the proximity between participants and objects effectively contributes to emotional engagement, communicative clarity, and persuasive effectiveness.

5. Audio Mode

The effectiveness of the audio mode was analyzed using van Leeuwen's (1999) framework, focusing on music, sound effects, silence, and tone. According to van Leeuwen (1999), auditory elements function as semiotic resources that contribute to meaning-making by

shaping emotional atmosphere, directing viewer attention, and reinforcing communicative intent. These semiotic resources were examined to determine their contribution to communicative clarity, emotional engagement, brand reinforcement, and the advertisement's persuasive goals.

Regarding the first aspect of the audio mode, the background music in the scene is classified as effective. The commercial features a cheerful and upbeat jingle with repetitive lyrics such as "Makuku, mamaku, aku sayang kamu" and "Makuku ku ku ku ku sayangku." According to van Leeuwen (1999), music functions as a semiotic resource that conveys emotional meaning, establishes mood, and enhances audience engagement. The positive rhythm and affectionate lyrics create a warm and joyful emotional atmosphere that aligns with the advertisement's family-centered theme.

The repetition of the brand name within the jingle strengthens brand reinforcement by increasing memorability and associating the brand with feelings of love and care. This auditory repetition enhances communicative clarity by clearly linking the emotional message to the brand identity. In addition, the cheerful and memorable melody supports persuasive effectiveness by making the advertisement more engaging and emotionally appealing. Therefore, the use of music effectively contributes to emotional engagement, brand reinforcement, and the overall persuasive impact of the commercial.

Regarding the second aspect of the audio mode, the sound effects in the scene are classified as ineffective. The commercial relies primarily on the continuous jingle and does not include additional sound effects, such as footsteps, environmental sounds, or other auditory cues that could enhance realism. According to van Leeuwen (1999), sound effects function as semiotic resources that provide contextual information and enhance the sensory realism of audiovisual communication.

The absence of distinct sound effects limits the auditory richness of the scene and reduces the potential for reinforcing the physical realism of the baby's movement. For example, the inclusion of subtle environmental sounds, such as footsteps or ambient household sounds, could enhance communicative clarity by strengthening the viewer's perception of the baby's interaction with the environment. Without these supporting auditory cues, the sound effects aspect does not significantly contribute to emotional engagement

or message reinforcement. Therefore, the sound effects are classified as ineffective.

Regarding the third aspect of the audio mode, the use of silence in the scene is classified as somewhat effective. The commercial primarily features continuous background music, with only minimal variation in sound intensity. According to van Leeuwen (1999), silence functions as a semiotic resource that can emphasize key moments, create contrast, and direct viewer attention to important visual or linguistic elements.

Although complete silence is not present, brief moments of reduced musical intensity allow viewers to focus more clearly on visual elements, such as the appearance of the brand logo. This variation supports communicative clarity by allowing important visual information to become more perceptually salient. However, the absence of more deliberate or pronounced use of silence limits its effectiveness as a semiotic strategy for emphasis. As a result, the use of silence contributes partially—but not fully—to communicative clarity and viewer attention, supporting its classification as somewhat effective.

Regarding the fourth aspect of the audio mode, the tone of the scene is classified as effective. The overall auditory tone conveyed through the jingle and vocal delivery is warm, affectionate, and emotionally reassuring. According to van Leeuwen (1999), tone functions as a semiotic resource that communicates interpersonal meaning and emotional orientation, shaping how viewers interpret the relationship between participants and the communicative message.

The affectionate tone reinforces emotional engagement by emphasizing the loving relationship between parent and child, which aligns with the advertisement's thematic focus on care and protection. This emotional association strengthens brand reinforcement by positioning the Makuku product as a symbol of parental love and responsibility. In addition, the comforting and reassuring tone enhances persuasive effectiveness by appealing to parental emotions and encouraging positive brand perception. Therefore, the tone effectively supports emotional engagement, brand reinforcement, and the communicative goals of the advertisement.

Overall, the multimodal analysis of Data 1 demonstrates that the commercial effectively integrates linguistic, visual, gestural, spatial, and audio modes to communicate its intended

message (Kress & van Leeuwen, 2001, 2006; van Leeuwen, 1999, 2006, 2011; van Leeuwen & Jewitt, 2001; McNeill, 1992). Based on the criteria of clarity of meaning, relevance to the communicative purpose, emotional engagement, and contribution to viewer comprehension, most modal aspects function effectively in reinforcing the portrayal of comfort, safety, and parental care associated with the Makuku product. The linguistic mode provides clear and emotionally meaningful verbal reinforcement, while the visual, gestural, and spatial modes guide viewer attention and strengthen emotional resonance through carefully structured representation and composition. The audio mode further enhances the emotional atmosphere, although its limited variation in sound effects slightly reduces its overall effectiveness. These findings indicate that the coordinated use of multimodal resources supports the commercial's persuasive intent and facilitates effective meaning-making, demonstrating how multimodal integration contributes to the overall communicative effectiveness of advertising.

Data 5



Picture 2. Mama Nikita Willy with a smiling expression, surrounded by

Context

This screenshot was taken at 6.8 seconds of the Makuku Diapers commercial video featuring Nikita Willy. On the screen, Nikita Willy, a well-known public figure, is sitting on the floor. She is surrounded by three children, one of whom is playing with a toy dinosaur, while the other two are approaching or interacting with the group. The background audio is: "Makuku, mamaku, aku sayang kamu, Makuku ku ku ku ku sayangku. Makuku, papaku, aku sayang kamu, Makuku ku ku ku sayangku. Makuku..."

Analysis

This data was analyzed using the multimodal framework proposed Kress & van Leeuwen (2001, 2006), van Leeuwen (1999, 2006, 2011), van Leeuwen & Jewitt (2001), and McNeill (1992). Each mode is evaluated based on its

ability to contribute to communicative clarity, emotional engagement, and reinforcement of the advertisement's central message, which is family love, care, and trust in the Makuku diaper product.

1. Linguistic Mode

The effectiveness of the linguistic mode was analyzed using the frameworks of Kress and van Leeuwen (2001) and van Leeuwen (2006), focusing on text structure, typography, and layout.

The spoken jingle, "Makuku, mamaku, aku sayang kamu... Makuku, papaku, aku sayang kamu..." plays a significant role in constructing emotional meaning. According to Kress and van Leeuwen, linguistic structure contributes to interpersonal meaning by shaping emotional relationships between participants and viewers. The repetition of the phrases "mamaku," "papaku," and "sayang kamu" emphasizes affection, reinforcing the association between the product and parental love. This repetition enhances memorability and emotional resonance, which are key indicators of effectiveness in advertising discourse.

The product images are placed in the lower-right corner, which allows them to remain visible without distracting from the central emotional interaction between Nikita Willy and the children. This balanced arrangement ensures that both emotional content and brand identity are communicated effectively without competing for attention. Because the layout successfully guides viewer attention while maintaining clarity and emphasis on the advertisement's core message, it is categorized as effective.

2. Visual Mode

The effectiveness of the visual mode was analyzed using the theories of Kress and van Leeuwen (2006), van Leeuwen (2011), and van Leeuwen and Jewitt (2001), examining colour, style, and perspective.

Color contributes to emotional and symbolic meaning. The dominant use of warm yellow, light blue, and neutral tones creates a comforting and welcoming atmosphere. Yellow is commonly associated with happiness and warmth, while light blue conveys calmness and trustworthiness. These color associations support the advertisement's emotional appeal and reinforce the product's association with care and safety.

According to van Leeuwen (2011), color functions as a semiotic resource that shapes emotional interpretation. In this scene, the harmonious color palette enhances emotional

engagement and aligns with the family-oriented message, making the color aspect effective.

Style reflects the social and cultural meaning conveyed through visual design. The modern, minimalist interior, characterized by clean lines, uncluttered space, and functional furniture, conveys order, comfort, and stability. This visual simplicity directs attention toward the interaction between Nikita Willy and the children rather than distracting viewers with excessive decorative elements.

This stylistic choice aligns with Kress and van Leeuwen's (2006) concept that visual style contributes to representational meaning by framing the social context. The clean and organized environment supports the perception of a safe and nurturing home, reinforcing the product's association with reliable childcare. Therefore, style is effective.

Perspective shapes viewer involvement and emotional proximity. The camera is positioned at a relatively low angle and close distance, creating a sense of intimacy between viewers and the family. According to van Leeuwen and Jewitt (2001), closer social distance increases emotional engagement and viewer identification with participants.

This perspective allows viewers to feel emotionally connected to the family interaction, strengthening the advertisement's interpersonal meaning. As a result, the perspective effectively enhances emotional engagement and is categorized as effective.

3. Gestural Mode

The effectiveness of the gestural mode was analyzed using the theory proposed by McNeill (1992), focusing on facial expressions, body language, and interpersonal interaction.

Facial expressions play a key role in conveying emotional meaning. Nikita Willy and the children display joyful and affectionate expressions, which communicate warmth, happiness, and emotional bonding. These visible expressions provide direct emotional cues to viewers.

According to McNeill (1992), facial expressions function as embodied semiotic resources that reinforce interpersonal meaning. The clear visibility of positive emotional expressions strengthens viewer engagement and supports the advertisement's emotional appeal. Therefore, facial expressions are effective.

Body posture and movement convey openness and emotional accessibility. Nikita Willy's

relaxed and open posture signals comfort and attentiveness, while the children's active and engaged movements convey playfulness and energy. These bodily actions reinforce the perception of a caring and supportive environment.

This aligns with McNeill's (1992) concept that gestures communicate meaning beyond spoken language. The body language effectively reinforces emotional warmth and product association with comfort, making it effective.

Interpersonal interaction contributes to relational meaning. The close and affectionate interaction between Nikita Willy and the children visually represents nurturing relationships and emotional closeness.

According to multimodal theory, such interaction strengthens interpersonal meaning by demonstrating social relationships. This interaction directly reinforces the advertisement's core message of family love and care, making it effective.

4. Spatial Mode

The effectiveness of the spatial mode was analyzed using Kress and van Leeuwen (2006) and van Leeuwen (2006), focusing on placement, directional flow, and proximity.

The family is positioned centrally, making them the primary focus of the scene. The product and logo are positioned peripherally, ensuring visibility without overshadowing the emotional interaction. This placement reflects van Leeuwen's (2006) principle of salience, where central elements attract the most attention. This supports both emotional engagement and brand recognition, making placement effective.

Viewer attention is guided naturally from the central family interaction toward the product and logo, creating a deliberate visual sequence. This structured flow encourages viewers to engage emotionally with the scene before recognizing the brand. According to Kress and van Leeuwen (2006), directional flow functions as a semiotic resource that guides viewer attention and structures the narrative progression of visual communication. In this context, the sequencing enhances narrative coherence and communicative clarity, demonstrating the effectiveness of directional flow.

The close physical proximity between Nikita Willy and the children visually represents emotional closeness and care. This spatial arrangement reinforces interpersonal meaning and strengthens the emotional resonance of the

scene. According to van Leeuwen (2006), proximity functions as a semiotic resource that conveys interpersonal relationships and social meaning by representing degrees of closeness or distance between participants and their environment. In this context, the deliberate spatial closeness effectively communicates warmth and relational intimacy. Therefore, proximity is effective.

5. Audio Mode

The effectiveness of the audio mode was analyzed using van Leeuwen's (1999) framework, examining music, sound effects, silence, and tone.

The cheerful jingle creates a positive emotional atmosphere and enhances memorability. According to van Leeuwen (1999), music functions as a semiotic resource that contributes to emotional framing and interpersonal meaning. The melody reinforces brand identity and emotional warmth. Therefore, music is effective.

The absence of additional sound effects reduces sensory richness and realism. Van Leeuwen (1999) explains that sound effects enhance environmental realism and support narrative immersion. Because the auditory experience relies solely on the jingle, sound effects are categorized as ineffective.

Brief moments of reduced musical intensity help highlight key visual elements such as the brand logo. Van Leeuwen (1999) argues that silence functions as a meaningful semiotic resource by creating emphasis and structuring attention. However, the continuous presence of music limits the strategic use of silence. Therefore, silence is somewhat effective.

The overall tone is warm, affectionate, and reassuring. According to van Leeuwen (1999), vocal and musical tone shape interpersonal relations and emotional orientation. This tone aligns with the advertisement's message and enhances emotional engagement. Therefore, tone is effective.

Overall, the multimodal elements in this scene work cohesively to construct a strong emotional and communicative message (Kress & van Leeuwen, 2001, 2006; van Leeuwen, 1999, 2006, 2011; van Leeuwen & Jewitt, 2001; McNeill, 1992). The visual, gestural, and spatial modes are effective in reinforcing themes of family love, emotional warmth, and trust, which are central to the advertisement's persuasive intent. The linguistic mode is somewhat effective, as the

spoken jingle enhances emotional engagement, while the on-screen linguistic elements have limited prominence. Similarly, the audio mode is somewhat effective: although the jingle and tone successfully create emotional appeal and brand memorability, the absence of sound effects and limited use of silence reduce the overall auditory depth. According to the multimodal framework, meaning emerges through the integration of semiotic resources, and in this scene, emotional interaction, supported by warm visual design and engaging sound, plays the dominant role in communicating the advertisement's core message and positioning the Makuku product as a symbol of care, comfort, and parental love

Data 10.



Picture 3. Baby Issa, being hugged by Mama Nikita Willy, is still wearing a Makuku diaper at 9 p.m

Context

This screenshot was taken at 16.0 seconds of the Makuku diapers commercial video featuring Nikita Willy. The screen shows Nikita Willy holding Baby Issa, with the clock showing 9 o'clock. The audio playing in the background is audible: "Makuku, mamaku, aku sayang kamu, Makuku ku ku ku ku sayangku. Makuku, papaku, aku sayang kamu, Makuku ku ku ku sayangku..."

Analysis

This data was analyzed using a multimodal social semiotic framework based on the theories of Kress & van Leeuwen (2001, 2006), van Leeuwen (1999, 2006, 2011), van Leeuwen & Jewitt (2001), and McNeill (1992).

1. Linguistic Mode

The effectiveness of the linguistic mode was analyzed using the frameworks of Kress and van Leeuwen (2001) and van Leeuwen (2006), focusing on text structure, typography, and layout.

Regarding the first aspect, the text structure is effective. The phrase "Disarankan untuk mengganti popok setiap 4 jam sekali" (Recommended to change diapers every 4 hours) is clear, concise, and directly related to the product's use and care instructions. According to Kress and van Leeuwen (2001), linguistic

structure contributes to meaning-making through the organization and sequencing of information. The use of numbers (1–12) arranged around the clock image emphasizes the concept of time, reinforcing the four-hour recommendation. This integration of verbal text and visual support enhances clarity and memorability. Therefore, the text structure is effective.

Regarding the second aspect, the typography is somewhat effective. Van Leeuwen (2006) explains that typography functions as a semiotic resource through font style, weight, and visual form, shaping interpersonal meaning and tone. The text uses a clear, rounded sans-serif font that appears friendly and approachable, aligning with the family-oriented nature of the product and contributing to a comforting and trustworthy atmosphere. However, the logo at the bottom right is relatively small, and the two Makuku diaper items beside it are undersized, making the product details difficult to read. Therefore, the typography is somewhat effective.

Regarding the third aspect, the layout is effective. According to Kress and van Leeuwen (2001), layout and placement influence information value and guide viewer attention. The message is positioned at the bottom of the screen, allowing viewers to first engage with the mother-child interaction before processing the informative text. The clock and numbers are centrally arranged around the mother and child, directing the viewer's gaze and emphasizing the time interval. This spatial arrangement clearly links the product to the timing suggestion. Therefore, the layout is effective.

2. Visual Mode

The effectiveness of the visual mode was analyzed using the theories of Kress and van Leeuwen (2006), van Leeuwen (2011), and van Leeuwen and Jewitt (2001), examining colour, style, and perspective.

Regarding the first aspect, colour is effective. According to van Leeuwen (2011), colour functions as a semiotic resource that conveys meaning through dimensions such as hue, saturation, brightness, and modulation. The dominant blue, yellow, and warm white tones create a comforting and nurturing atmosphere. Yellow evokes happiness and optimism, supporting the family-centered theme, while blue reinforces calmness and trust. The white product packaging stands out against the vibrant background, symbolizing cleanliness and quality. Therefore, the colour is effective.

Regarding the second aspect, style is effective. Van Leeuwen and Jewitt (2001) explain that visual style reflects social values, identities, and cultural meanings embedded in visual representation. The modern, minimalist room design emphasizes comfort and practicality, creating a welcoming and family-friendly environment. The uncluttered arrangement allows natural interaction, while the product packaging is subtly integrated into the scene. Therefore, the style is effective.

Regarding the third aspect, perspective is effective. According to Kress and van Leeuwen (2006), perspective, angle, and framing influence viewer involvement and relational positioning within visual communication. The eye-level angle creates a natural and relatable interaction, allowing viewers to feel connected to the family moment. The slightly angled shot enhances openness and directs attention toward the product packaging. Therefore, the perspective is effective.

3. Gestural Mode

The effectiveness of the gestural mode was analyzed using McNeill's (1992) theory, focusing on facial expressions, body language, and interpersonal interaction.

Regarding the first aspect, facial expression is effective. According to McNeill (1992), gestures and facial expressions function as meaningful bodily signs that convey emotional states and communicative intent. Nikita Willy's joyful expression communicates warmth, happiness, and affection toward Baby Issa, while the child's relaxed demeanor reinforces the sense of comfort and security associated with wearing Makuku Dry Care diapers. Therefore, facial expression is effective.

Regarding the second aspect, body language is effective. McNeill (1992) explains that bodily movements and posture contribute to interpersonal meaning and relational positioning. Nikita Willy's welcoming posture, with her arms around the child, highlights her nurturing role and strong emotional bond. Baby Issa's relaxed body language further emphasizes comfort and care. Therefore, body language is effective.

Regarding the third aspect, interpersonal interaction is effective. In McNeill's (1992) framework, gesture operates within interactional contexts to construct shared meaning between participants. The intimate and nurturing exchange between Nikita Willy and Baby Issa conveys trust and affection, reinforcing the message that Makuku products support moments of security

and emotional connection. Therefore, interpersonal interaction is effective.

4. Spatial Mode

The effectiveness of the spatial mode was analyzed using Kress and van Leeuwen (2006) and van Leeuwen (2006), focusing on placement, directional flow, and proximity.

Regarding the first aspect, placement is effective. According to Kress and van Leeuwen (2006), spatial placement determines information value and salience within a composition. The mother and child are centrally positioned, making them the primary focus and reinforcing their emotional connection. The product packaging is placed subtly in the bottom right corner, ensuring visibility without distracting from the main interaction. This balanced distribution of elements supports both emotional emphasis and brand presence. Therefore, the placement is effective.

Regarding the second aspect, directional flow is effective. Kress and van Leeuwen (2006) explain that compositional structures guide viewer attention and organize visual sequences. The Makuku logo at the top left attracts initial attention, establishing brand identity. The centrally positioned clock graphic then directs focus to the functional message of changing the diaper every four hours. The mother-child interaction follows, strengthening emotional engagement, and finally the product packaging appears at the bottom right as a visual reminder. This structured progression (from brand, to function, to emotion, to product) creates a coherent and persuasive visual narrative. Therefore, the directional flow is effective.

Regarding the third aspect, proximity is effective. Van Leeuwen (2006) argues that spatial distance functions as a semiotic resource that conveys interpersonal relationships and social meaning. The close physical proximity between Nikita Willy and Baby Issa highlights their emotional bond and sense of security. The product packaging is positioned at a slight distance to maintain visibility while preserving the intimacy of the interaction. The softly blurred background elements further maintain focus on the central relationship while suggesting a safe and comfortable environment. Therefore, the proximity is effective.

5. Audio Mode

The effectiveness of the audio mode was analyzed using van Leeuwen's (1999)

framework, examining music, sound effects, silence, and tone.

Regarding the first aspect, music is effective. According to van Leeuwen (1999), music functions as a semiotic resource that shapes emotional atmosphere, interpersonal meaning, and memorability. The background jingle with the lyrics, “Makuku, mamaku, aku sayang kamu, Makuku ku ku ku ku sayangku. Makuku, papaku, aku sayang kamu, Makuku ku ku ku sayangku...,” is catchy and repetitive, creating an engaging and memorable audio experience. Its rhythm and repetition reinforce brand recall and emotional warmth. Therefore, the music is effective.

Regarding the second aspect, sound effects are ineffective. Van Leeuwen (1999) explains that sound effects contribute to realism, immersion, and environmental context in audiovisual texts. In this commercial, only the jingle is heard from beginning to end, without additional ambient or contextual sounds. The absence of varied sound effects reduces sensory richness and limits narrative depth. Therefore, the sound effects are ineffective.

Regarding the third aspect, silence is somewhat effective. According to van Leeuwen (1999), silence functions as a meaningful semiotic resource that creates emphasis and structures attention. Although silence is used sparingly, brief reductions in musical intensity help highlight key moments, such as the appearance of the “Makuku” brand name. However, the continuous presence of the song limits the strategic impact of silence. Therefore, silence is somewhat effective.

Regarding the fourth aspect, tone is effective. Van Leeuwen (1999) notes that tone shapes interpersonal relations and emotional orientation in audiovisual communication. The overall tone of the scene is warm, affectionate, and reassuring, emphasizing the close relationship between mother and child. This comforting tone resonates with parents who wish to provide the best for their children. Therefore, the tone is effective.

Overall, the multimodal elements in this scene work together effectively to communicate both the functional and emotional meanings of the advertisement, although not all modes contribute with equal strength (Kress & van Leeuwen, 2001, 2006; van Leeuwen, 1999, 2006, 2011; van Leeuwen & Jewitt, 2001; McNeill, 1992). The linguistic mode is somewhat effective because the instructional text, “Disarankan untuk mengganti

popok setiap 4 jam sekali,” supported by the clock graphic, clearly conveys product usage guidance and constructs ideational meaning through structured information design. However, the reduced salience and smaller size of the product logo and packaging text limit their visibility, thereby weakening their contribution to brand reinforcement. The visual mode is effective, as the warm color palette, minimalist style, and eye-level perspective create a nurturing and trustworthy environment. Colour operates as a semiotic resource shaping emotional meaning, while style and perspective structure viewer involvement and social meaning. Together, these elements symbolically associate the product with care, comfort, and reliability. The gestural mode effectively conveys interpersonal meaning through Nikita Willy’s affectionate facial expression, protective posture, and intimate interaction with Baby Issa, visually representing emotional bonding, safety, and comfort. Similarly, the spatial mode is effective, as the central placement of the mother and child, supported by the clock graphic and strategically positioned product packaging, guides viewer attention and organizes compositional meaning. Proximity further constructs relational meaning by visually representing intimacy and distance. In contrast, the audio mode is somewhat effective. Although the repetitive jingle and warm tone enhance emotional engagement and brand memorability, the absence of varied sound effects and minimal strategic use of silence reduce auditory emphasis and semiotic richness. Taken together, the integration of linguistic, visual, gestural, spatial, and audio modes demonstrates how multimodal semiotic resources collectively position the Makuku diaper as both a functionally reliable product and an emotionally supportive element within a caring parent–child relationship.

CONCLUSION

This study examined the multimodal construction of meaning in the Makuku diapers commercial featuring Nikita Willy using the multimodal social semiotic framework of Kress and van Leeuwen and van Leeuwen. The findings show that the advertisement effectively integrates linguistic, visual, gestural, spatial, and audio modes to communicate emotional meaning, reinforce brand identity, and enhance persuasive impact, although the effectiveness varies across specific aspects.

The linguistic mode is generally somewhat effective. The spoken jingle strengthens

emotional engagement and brand recall through repetition and affectionate language, consistent with the interpersonal function of language in multimodal communication. However, the reduced salience of smaller logos and product text limits the overall effectiveness of typography and written linguistic elements.

The visual mode is consistently effective. Color, style, and perspective function as powerful semiotic resources that communicate warmth, comfort, and trust while creating emotional closeness between viewers and participants. These visual elements align with multimodal theory, which emphasizes the role of visual composition in shaping emotional and symbolic meaning.

The gestural mode is overall effective, although facial expressions are not always visible. In scenes where facial expressions are clearly shown, they strengthen emotional connection. In scenes where they are not visible, body language and interaction compensate by effectively conveying comfort, affection, and relational meaning. This demonstrates that gestural meaning can be communicated through multiple embodied resources, not facial expression alone.

The spatial mode is effective in guiding viewer attention and organizing meaning. The placement, directional flow, and proximity of elements follow compositional principles that emphasize emotional interaction while maintaining brand visibility. These spatial arrangements support narrative clarity and reinforce emotional engagement.

The audio mode is generally effective. The jingle and tone enhance emotional appeal and brand memorability, functioning as affective semiotic resources. However, the absence of varied sound effects and the limited use of silence reduce the overall richness of the auditory experience.

Overall, the commercial demonstrates effective multimodal integration, particularly in visual, gestural, spatial, and core audio aspects, while linguistic and certain audio aspects are somewhat effective due to reduced salience and variation. These findings support multimodal social semiotic theory, which emphasizes that advertising meaning emerges from the interaction of multiple semiotic modes working together to create persuasive and emotionally engaging communication.

Based on these findings, several recommendations can be proposed. First, advertisers should enhance the visibility and salience of written linguistic elements to strengthen message clarity and brand recognition. Second, incorporating varied sound effects and more strategic use of silence could enrich the audio mode and improve emotional impact and realism. Third, future advertisements should ensure greater visibility of facial expressions to maximize emotional clarity and interpersonal engagement. Finally, future research should examine larger datasets, different advertising contexts, or audience reception to further explore how multimodal resources influence advertising effectiveness.

REFERENCES

- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis—A problem-oriented introduction*. De Gruyter Mouton.
- Chandler, D. (2007). *Semiotics: The basics* (2nd ed.). Routledge.
- Danesi, M. (2021). *The semiotics of everyday life*. Bloomsbury Academic.
- Davis, S. (2019). Advertising's emotional appeal and its impact on consumer behavior. *Marketing Theory*, 24(4), 482–495. <https://doi.org/10.1016/j.marktheo.2019.06.003>
- Fairclough, N. (1992). *Discourse and social change*. Polity Press.
- Ferguson, R. (2024). Emotional marketing: Connecting with the modern consumer. *Journal of Advertising*, 53(2), 112–127. <https://doi.org/10.1080/00913367.2024.1765902>
- Global Brands Magazine. (2024). The power of empathy in brand advertising. <https://www.globalbrandsmagazine.com/empathy-in-advertising/>
- Halliday, M. A. K. (2004). *An introduction to functional grammar* (3rd ed.). Arnold.
- Hughes, C. (2024). Commercial video production: An overview. *Journal of Media Production*, 11(3), 45–59