

## Voice and Visual Framing in Television News: Influence of Narration on Viewer Perception in Kazakhstan

Kydyr-Zhan Darynuly Mukhatayev  
Independent University "Turan"  
Kazakhstan

### ABSTRACT

The paper examines the alliance of voice narration in concert with the visual framing of televised news and how that relationship contributes to the perceptions of audiences who view televised news in the post-Soviet media space of Kazakhstan, which has experienced both the effects of state influence and media pluralism. Reflecting the elements of framing theory, narrative paradigm, visual semiotics as well as the aspects of receiving theory, the study establishes the influence of frames and codes of tones, pace, and the language used in voice-overs together with symbolic images and visual patterns over the trust of audience, which affects the emotional attachment and political sense. Using the convergent approach to a mixed-method study, data were analyzed based on 30 news items of large Kazakhstani broadcasters and focus groups were completed with a focus group made up of 24 interviewees of all age groups, languages, and geographical backgrounds. The results indicate the predominance of authoritative male voice and nationalistic imagery in entrenching state legitimacy and social harmony discourse. Although the target audiences, including older, rural, and Kazakh-speaking viewers, generally accepted these frames, younger, urban, and Russian-speaking readers were less compliant with them and frequently negotiated or challenged the messages that were supposed to be perceived. Such outcomes indicate the context-sensitive, two-fold character of media framing and acceptance in transitional democracies, as it may be associated with media ethics, literacy, and stance on the state-controlled news context.

**Keywords:** Voice-over, Visual Framing, Media Narratives, Television News, Kazakhstan, Viewer Perception, Media Influence

### 1. INTRODUCTION

The power of television news has remained a prevalent source of information among worldwide audiences, even in the era of digital media. Its powerful impact on society has been primarily attributed to its capacity to ensure a confluence of visual and audio activities in generating interesting stories that influence societal views regarding issues of the day (Grabe & Bucy, 2009). Specifically, contextualizing the news using pictures and voice-over comments is crucial in how people receive reality, judge its reliability, and make judgments (Entman, 1993; Iyengar, 1991). When dealing with a speech, we often make it much weaker or stronger with the help of the tone, the tempo, and the emotionalism of the voice, and this is the way how the visual manipulations in the broadcast can be supported or beaten. So, we direct the viewer without telling it straight (Esser, 2008).

Voice and visual framing, however, involve stylistic features and are components of media narratives that tend to parallel or support specific ideological, political, or cultural interests (Hall, 1997). Although there has been considerable research on such dynamics in the Western world, little research has been dedicated to studying how such framing techniques appear in a post-Soviet state, which is characterized by the impact of state authorities and the growing privatization of the media environment, e.g., in the country of Kazakhstan

(Nurbekova, 2018). Television is the predominantly circulating form of news in Kazakhstan, particularly with elderly communities and regardless of location, so it is a significant medium of influence on the civilian discourse (Open Society Foundations, 2021).

The methods employed in the narration of television news and in editing the television picture in the context of the development of the media in Kazakhstan, where regulation systems and journalist ethics are in a state of forming, might have other purposes than information. They can also quietly enforce the imaginaries of the dominant, favor the national identity, or deflate the voice of diss criticism (Junisbai, 2010). This concerns the ethical aspects of media framing, its implication to the populace, and its application to democracy. In addition, local consumers do not interpret these messages similarly along the cultural, linguistic, or age lines, which has formed a web of media reception worth academic research.

This paper fills this gap by examining how voice narration and visual framing affect news reports on Kazakhstan television. Particularly, it addresses the issue of the influence of the combination of spoken commentary and visual images on the perception of the viewer, their trust, and emotional reaction. This study is based on both framing theory (Entman, 1993) and narrative theory (Fisher, 1984), but more specifically on semiotic and reception-based media studies. Concentrating on the case of Kazakhstan, the study, in addition to transferring a more globally inclusive picture of how media effects are conceptualized and carried out, throws some light on the peculiarities of how audiovisual rhetoric works in the context of transitional media conditions.

The following research questions guide the study:

1. What role do voice Narration and visual pictures play each time framing the news in Kazakhstani television?
2. How do these factors affect the viewer's perception, especially regarding credibility, emotional appeal, and political judgment?
3. Do the viewers have culturally defined response patterns to various combinations of voice/visuals?

This study will provide theoretical and empirical contributions to the bodies of knowledge of media and communication studies by adopting a mixed-methods design that involves content analysis and audience responses collected using a qualitative approach to research.

## 2. LITERATURE REVIEW

This body of knowledge regarding television news framing is complex and multidimensional, incorporating theoretical, empirical, and cultural studies approaches. To situate this research in the proper context of academic studies, this review examines six primary strands of literature: a framing theory, voice narration in the news, visual veiling approaches to framing, the relationship between voice and visuals, the theory of audience reception, and the media environment in Kazakhstan.

### 2.1 Framing Theory and Television

The framing theory can still be considered one of the paradigmatic theories in media and communication, focusing on the role of media in shaping understanding among the population. In a study by Entman (1993), framing was defined as choosing some aspects of reality and highlighting them to encourage specific meanings. Frames put problems in place, diagnose the

issues, reach moral conclusions, and give solutions. This theoretical framework is highly applicable to studying the contents of news broadcasts, in which the judgment of the editor regarding the language usage, structure, and delivery pattern can change the collective thoughts to a great extent.

Framing in television news does not refer only to oral framing. The convergence of visual and sound features further extends the framing capabilities of the television so that the relayed messages can be conducted in a more covert and immersive way, as Grabe and Bucy (2009) point out. Each time a TV producer competes for an ideological signal through the designing of the broadcasts, e.g., through the arrangement of stories, use of background music, and texting on the screen, all of the above have an accumulative effect on how the audience interprets problems. Further, Iyengar (1991) distinguishes episodic and thematic framing, pointing to the idea that episodic framing individualizes issues (e.g., presents a sole protester). In contrast, the latter frames issues in perspective, i.e., present the long-lasting causes (e.g., a protester). Time limits and the desire to attract and capture the audience usually result in TV news adhering to episodic framing by not considering the structural factors of the event.

This framing power is mighty in media systems that have some regulations or political influence within the media. In this case, frames may not come out instantly but instead be influenced by institutional agendas. Voltmer (2013) highlights that frames in transitional or authoritarian media systems support ideologies of dominance and counteract any opposition by opportunistically reporting stories in favor of some rather than others.

## 2.2 Voice Narration as Framing Device

Voice narration in television news is both a descriptive and a framing system. It provides a tone of emotion, context of interpretation, and even a form of morality to the visual material on display. Bucy and Newhagen (1999) affirm that the human voice conveys paralinguistic information in tonal, amplitude, rhythmic, and inflectional messages, which can graciously influence viewers' opinions. As an illustration, a positive and aggressive tone can add value to a statement, and a timid or dull tone can undermine its credibility.

Voice-over narration has been investigated in terms of both cognitive and emotional-presenting framing. According to Geise and Baden (2014), narration can also focus viewers' attention on some news story details as a semantic bridge between the unaided footage and its implication. This is especially important when the image is ambiguous or emotionally connected, and the narrator's voice can help eliminate interpretive uncertainty. It is impossible to anticipate how narration will be equal or unequal to the subjects in emotionally sensitive settings, like natural disasters, political demonstrations, or aggressive conflicts.

In the post-Soviet states, the voice of the voice-down tends to copy the methods of the Soviet time, still dominating and sounding in the voice of the critical, authoritative man (Schimpfossl & Yablokov, 2020). These stylistic decisions might not be impartial because they arouse the feeling of institutional power and authority. Reporting in Kazakhstan uses an ideology of national pride or concern that guides the readership to be compliant or loyal in their views. This aligns with the opinions of Esser (2008) that political news cultures are connected to unique chains of stylistic traditions that determine how stories are told and heard.

## 2.3 Visual Framing in News Media

The perceptual part of news on TV can be incorporated into potent emotional and ideological chunks. In contrast to spoken or written language, visuals are basically in the realm

of non-verbal communication, and they primarily ride around critical thinking to generate emotional reactions. According to Messaris and Abraham (2001), images also serve implicit argumentative purposes, i.e., what is covered, covered, and foregrounded can create evaluative meaning without actually telling it. For example, a voice-over can be neutral or ambiguous, but long crying scenes of children or dramatic shots of police or police-involved struggles will support one type of narration.

The visual framing covers a variety of devices: framing shots, editing sequences, color schemes, the usage of archival material, and on-screen placement of text. All the options can enhance or negate the contextual text. For example, camera angles often give power and authority to the object in the scene when the subject exhibits an angle upwards; in other instances, the angle may indicate vulnerability when the camera angle goes downwards. The fast cutting can make the whole thing chaotic, and the slow motion can make something romantic or dramatic (Coleman, 2010).

The Kazakhstani blends the Russian and Western aesthetics of the TV console. Still, the case de la spectrum often uses visual strategies to reinforce the patriotic perspective or social unification. National monuments or state officials interacting with people or young people in traditional attire are typical depictions that provoke unity and stability (Nurbekova, 2018). Together with properly constructed narration, these visual signals put the concept of national identity in contexts that the state police and the desperate anticipation of the audience find congruous.

#### **2.4 Interaction Between Voice and Visuals**

Voices and visuals have the power of framing, but they may have more power than the sum of these two parts. The combination of what is discussed and what is demonstrated may either support or be at variance with the message. According to the arguments by Knobloch-Westerwick and Hastall (2006), the employment of multimodal consistency (a situation in which voice and visual information match) improves memory storage and interpretational transparency. Conversely, suppose there is no correlation between the narration and visuals. In that case, the audience might feel cognitive dissonance when they doubt the credibility or purpose of the broadcast.

In situations with political messaging, this dissonance might not necessarily be an accident. According to Kim and Barnett (1996), visual and verbal contradictions constitute one tactic the state-controlled media may employ in sending mixed messages: the pictures will indicate unrest, but the narration will remind the audience that everything is stable and controlled. Such methods can cover politically sensitive issues such as protests, corruption activities, or foreign affairs in Kazakhstan. Through a news report and narration in a tone that minimizes the importance of the case or places blame on outside players, it is possible to manipulate perception without blatant censorship, as may be shown in the demonstration video.

Researchers have equally stated, and it is worth highlighting, that the priority of voice and visual introduction can dictate narrative salience. Grabe and Bucy (2009) present that visuals and voice as leading media produce a different impression on the consciousness than narration that precedes the image. This time ordering is essential in how viewers receive and internalize information, particularly in fast-moving television programs.

## 2.5 Viewer Perception and Reception Theory

Media interpretation is not a passive process determined by the audience; it is a social, cultural, and psychological process of interpretation. The encoding/decoding model offered by Stuart Hall (1980) can be seen as a crude structure seen through the dynamics. Whereas producers of media materials encode meanings in messages, receivers of the messages are seen to decode them with meanings, which may represent the intended meanings, negotiate such, or work against it. This implies that the news item can have different interpretations per demographical groups based on culture, political orientation, and media literacy.

This decoding process is most severe and complicated in Kazakhstan, where people in a nation with generational differences and diverse access to media enjoy linguistic diversity. Different generations who grew up under the Soviet regime might perceive formal voice narration and traditional imagery patterns better than young audiences who think them old-fashioned or propagandistic (Kudaibergenova, 2016). Besides, the narrative language Kazakh, or Russian, has specific political and cultural implications that influence the sense. Broadcasts in Russian are more objective or purpose of modernity. In contrast, Kazakh language broadcasts focus on national identity and tradition.

In addition, there is the legacy of state-controlled media in Kazakhstan. It plays to the overall skeptical nature of viewers who can watch television news as skeptical or resigned viewers. According to a study by Beissinger (2013), in post-Soviet countries, political symbolism can be projected according to a twofold way of understanding, where the official messages are received in both the surface and subversive manner. These two perceptions create even more complexity in voice reception and visual framing.

## 2.6 Television News in Kazakhstan

Kazakhstan is a special study of media framing because it is a hybrid media environment structure. Formally democratic and, in a real sense, authoritarian, the country provides a level of media pluralism and a firm grip on mainstream TV. According to Freedom House (2022), the media situation in Kazakhstan is considered "Not Free" because the government and the country own the biggest channels and have repressive press laws. However, TV scores the highest in impact, especially in rural environments where the net connection might be unavailable.

According to Nurbekova (2018), Kazakhstani television lies in the paradigm of state discourse: Though the form of the TV is clear of modern journalistic standards, which include professional production, formatting that follows international formats, and various programming approaches, the content itself acts as a state propaganda. It is not exacted using flagrant propaganda but concealed methods like pick-and-choose, manipulative narration, and heart-touching imagery. For example, the media coverage of the president's affairs is commonly made with soft light, inspirational music, and favorable commentary, producing the same impression that the president is a benevolent leader.

Voice narration of the Kazakhstani news is strongly connected with the formation of national identity. The issue of respect and legacy is often indicated in the narration of state ceremonies or anniversaries of historical events, as narrators use tones of a solemn, reverent voice. Images of cultural images, flags, and military parades also support patriotic appeal. Even these framing techniques represent a larger ideological agenda to stimulate national solidarity while downplaying social upheavals or disgruntled political attitudes.

## 2.7 Synthesis and Research Gap

Even though voice or visual framing has been successfully analyzed in varying contexts, little research was done investigating the combined influence of both voice and visual framing on audience perception in the post-Soviet media environment. In addition, the literature on Kazakhstani television is limited, particularly in its ability to explain the role of audiovisual methods in viewer interpretation and political socialization. To help address this gap, this paper examines the interaction between voice narration and frame-breaking images and how they are used to shape the opinion of the populous in Kazakhstan, thus further validating current models on media framing in a culturally unique and non-western society.

## 3. THEORETICAL FRAMEWORK

The issue of television news narrative and imagery affecting the audience's perception in Kazakhstan implies a multidimensional theoretical methodology. Considering the multifaceted nature of mediated communication, especially in transitional media systems, this paper relies upon four interrelated theories: the Framing Theory, Narrative Paradigm Theory, Visual Semiotics, and Reception Theory. These views provide the analytical depth to analyze meanings' construction, coding, and interpretation in media formats such as televised news.

### 3.1 Framing Theory

The Framing Theory offers the core prism by which this paper focuses on the nature by which television news frames and represents reality. Frames, as described by Goffman (1974), are a schema of interpretation that provides a person with a way of understanding a social situation. This idea was polished by Entman (1993), who presented the definition of the notion as follows: to frame is to select some aspects of the perceived reality and make them more salient and to advance a particular way of defining a problem and diagnosing its causes, moral considerations, and proposed solutions (p. 52).

There are several ways of framing the television news: via verbal script, visual composition, editing, tone, and even sequence. Of most concern, especially, is multimodal framing, which involves the compounding of narrative and images to establish or confirm boundaries of interpretation. D-Angelo and Kuypers (2010) explain that frames promote journalistic decisions and institute ideologies, thus, imperative instruments in establishing political and cultural discourses.

Framing theory is of particular relevance to the semi-authoritarian Kazakhstan media environment. It has been demonstrated that state-aligned broadcasters apply a similar framing pattern to achieve national unity, authorize power, and isolate opposition (Nurbekova, 2018). This paper goes a step further to apply framing theory on the use of audio and visual elements to create frames in voice narration and imagery sequencing, camera angles, and symbols that influence the cognitions and affections of the viewers.

### 3.2 Narrative Paradigm Theory

This study will use the Narrative Paradigm Theory, which is advanced by Fisher (1984), to gain clearer insight concerning the arguments of the persuasive power of voice-over narration. Fisher claims that human beings are storytellers by nature, and we are more likely to find ourselves convinced by narratives that mirror our truths of life and values than logical reasoning. Based on this theory, every human communication could be measured using the coherence and fidelity of the stories they convey.

Though often intended to be objective, television news is usually told within narrative structures, which makes the stories more identifiable and emotionally gripping. These stories involve characters (heroes/villains), conflict, resolution, and the presence of moral messaging. Voice narration is central in building these narratives because it provides continuity, interpretation, and moral indications (Chouliaraki, 2006).

The situation in the Kazakhstani media environment, where manifestations of national existence and stability in achieving development are the constant manifestations of the state ideology, allows narrative voice-over to reinforce ideological discourses effectively. Political leaders are usually depicted in news segments as saviors or change agents. In contrast, the dissension or crisis is discussed as the outside or aberrant. The Narrative Paradigm Theory is used in this study to evaluate the role of voice narration that serves beyond the scope of communicating the news but actually to story-tell the news in a way that complies with state-approved values.

### 3.3 Visual Semiotics

Verbal narration often structures the what of a news story, whereas visual elements usually define the how. On this basis, Visual Semiotics, the study of signs and meaning in pictures, is used to interpret how television images convey the ideology. Barthes (1977) distinguishes between the denotative (or literal) and connotative (or cultural/symbolic) meanings of images and states that images are never neutral. All the visual decisions, including the spectrum of colors and camera positioning, may help to direct observers to some specific interpretations gently.

In TV news, visual semiotics are used to unravel the hidden, symbolic narration of the visuals selected to go with the narration. For example, repeated national symbols (tables with flags, national costumes, and national monuments in the street) may express patriotism and legitimacy. In contrast, chaotic scenes (anti-governmental demonstrations, fire, empty shops in the supermarket) may create an impression of disorder or danger. These connotations may also be strengthened by effectively collating such pictures into conciliatory voice-overs (Messaris & Abraham, 2001).

Due to the popularity of visual representation alongside ceremony or political actions, Kazakhstan may decline such representational manner of visibility. Semiotic study in the sphere of visual other codes discloses the way that state stories are visually 'inscribed into the daily mass media. The present research uses the concepts of the semiotics of imagery to locate the emergent patterns of visual similarities and themes and explain how they combine with the edited word to produce seamless imminent arguments.

### 3.4 Reception Theory

Although understanding how messages are created can assist us through the use of framing, narrative, and semiotic theories, the Reception Theory is an approach that draws attention to the characteristics of message reception. Hall (1980) developed his encoding/decoding model, in which media texts are not consumed by the audience passively but are actively identified by those using them according to their cultural location, social background, and experiences.

Hall identifies three possible decoding positions:

- **Dominant-hegemonic:** The viewer accepts the intended meaning.
- **Negotiated:** The viewer partly agrees but resists some aspects.

- **Oppositional:** The viewer rejects the intended meaning entirely.

This model is mainly practicable in the case of Kazakhstan, where the audiences broadly differ in age, language, education, and the availability of alternative media sources. For example, a Kazakh-speaking viewer in a rural area would understand the news story with a governmental frame differently from an urban Russian-speaking young adult exposed to the international press. In addition, the Soviet past of media control in Kazakhstan means a culture of media suspicion and multi-level interpretation (Beissinger, 2013). Viewers can perceive a television report's approved and alternative meanings.

This study considers the complexity of viewer interpretation by introducing the Reception Theory, which does not presume a generalized effect of narration and imagery. Instead, it focuses on how various audience members negotiate the meaning of television news, especially regarding the combination of voice and pictures.

### Theoretical Integration

Together, these four frameworks provide a robust foundation for this study:

- **The Framing Theory** reveals the way voice and pictorial meaning build interpretations.
- **The Narrative Paradigm Theory** states that voice narration captivates the audience's perception and emotional connection with the story.
- **Visual Semiotics** exposes the power of images in strengthening meaning.
- **Reception Theory** acknowledges that audience interpretation is diverse due to the context and identity.

This embedded theoretical concept enables an in-depth excursion into how, why, and with what consequences in Kazakhstan, television news, through voice and visibility framing, influences viewers' perceptions.

## 4. METHODOLOGY

This study uses a mixed-method research design to examine how voice narration and visual framing affect viewer perception during Kazakhstani television news. The proposed method to be used would combine qualitative content analysis of televised news broadcasts with focus group discussions to shed light on the myriad of ways in which framing strategies are incorporated into the production of televised news as well as to understand the experiences of viewers who have been exposed to such content. The mixed-methods design is especially effective in examining media interrelations, as it makes it possible to achieve the categories triangulation and a better insight into the intersections of the media content and its interpretation (Creswell & Plano Clark, 2018).

### 4.1 Research Design

The current study used a convergent mixed-method design to examine the effects of voice narration and visual framing on the perception of televised news by the audience in Kazakhstan. Considering the multimodality of television products, content analysis and focus group discussions were used to allow the research to investigate how media frames are created and how the different viewers make meanings of these media frames. Convergent designs can be effectively used in media research, as media studies need a wide range of information to be collected simultaneously and analyzed to produce a broad scope of information (Creswell & Plano Clark, 2018).

#### 4.2 Sampling Techniques and Data Sources

The procedures for identifying the media content and sampling the participants were based on purposive sampling techniques. For the content analysis, three prime-time news segments of 30 total were selected for three months (January 2025, March 2025, and March 2025) in Kazakhstan's three largest television broadcasting companies: Khabar TV, Kazakhstan TV, and Channel 31. They elected these broadcasters because of their wide national outreach and varying degrees of editorial freedom (Nurbekova, 2018). The segments were chosen based on politically or socially important issues like national events, economic development, civil unrest, and the population's health.

Meanwhile, a focus group of 24 people was assembled with four participants in four locations (Almaty, Nur-Sultan, Shymkent, and Semey) to provide age, gender, education, language preference, and urban-rural residency representation. Such demographic diversity also explains Kazakhstani media audiences' cultural and lingual diversity (Kudaibergenova, 2016).

#### 4.3 Data Collection Methods

Each previously selected news segment was transcribed and followed with verbal and visual framing analysis. Referring to the classic paradigm (Entman, 1993; Grabe & Bucy, 2009), the approach was to check on the narration level, method of delivery, camera angle, imagery, and emotion suggestiveness. A hybrid coding strategy was resorted to combining deductive and inductive codes inspired by the data. NVivo 14 was used to analyze transcriptions, and intercoder reliability was also determined by calculating the Cohen, which was as high as 0.80, proving that the disagreement code did not reach a significant critical mark (Landis & Koch, 1977).

Four focus groups with six respondents each were established for the audience study in four selected locations. According to the participants' preferences, the sessions took place either in Kazakh (or Russian). They were estimated to occupy about 60 to 90 minutes. In each session, the participants watched previously selected news clips. They discussed them in semi-structured interviews with prompts on credibility, emotional tone, symbolic imagery, and trust in media. Such a structure allowed communication and shared meaning, deepening the level of insight (Krueger & Casey, 2015).

#### 4.4 Data Analysis

The theoretical models that informed content analysis were Welcome to Our PLANET (Framing Theory, Entman, 1993), Visual Semiotics (Barthes, 1977), and Narrative Paradigm Theory (Fisher, 1984). The codes' systems were put into the categories of higher levels of narration: authoritative, national symbolism, emotive voice modulation, and visual-verbal synchronization. A thematic matrix was formulated to trace patterns of the 30 news segments.

Thematic analysis of the transcripts of the focus groups was carried out by applying the six-step model developed by Braun and Clarke (2006). The emergent themes were the state power understood by narrating, emotional packaging, mistrust of imagery, and disparity of media credibility between generations. These data have also been presented in Stuart Hall's (1980) encoding/decoding model, and the viewers have shown to be dominant, negotiated, and oppositional readings as per their background and media literacy. Table 1 describes coding categories created according to the available literature (Entman, 1993; Messaris & Abraham, 2001).

**Table 1: Coding Dimensions and Indicators Used in Content Analysis**

<b>Framing Dimension</b>	<b>Indicator</b>	<b>Description</b>
Voice Tone	Authoritative / Neutral	Institutional authority or objectivity was expressed through pitch, tempo, and assertiveness.
Visual Symbolism	High / Moderate / Absent	The presence of flags, national icons, leaders
Synchronization	Congruent / Incongruent	Voice-over narration / Aural effect degree of coherence with visual imagery.
Emotional Appeal	High / Medium / Low	Emotional tone guided by voice and visuals

#### **4.5 Ethical Consideration**

Informed consent was obtained in writing before all participants took part. The research provided for confidentiality, anonymity, and the voluntary nature of participation as required by the ethics of social research. Due to the politically sensitive nature of the responses, the conversations took place in a non-coercive atmosphere, and they were made without exerting any pressure to assimilate to any prevailing narratives. News content published publicly was analyzed, which adhered to the ethical standards of data use.

#### **4.6 Limitations of the Methodology**

Although the mixed-method study contributed to a rich data set and a triangulation of results, there are limitations associated with the study. Purposive sampling of news segments is only valid in context, but this restricts generalizability. Moreover, the social desirability factor was introduced due to focus group environments, especially when addressing hot political topics (Beissinger, 2013). Finally, the cross-sectional design does not allow measuring the effects of framing in the long term. However, the methodology is a strong way to comprehend the framing and audience perception strategy in the shifting news media.

### **5. RESULTS**

The report will include the results of the two main data sets that have been collected: (1) a content study of 30 television news stories in three large Kazakhstani television outlets and (2) focus group interviews with 24 diverse participants in terms of age, gender, and nationality. Its analysis shows a few prevailing patterns regarding the use of voice, visual framing, and the significant difference between the interpretation of the frames by the viewers. Their results are described in two sections: framing in media content and audience perceptions, structured thematically in the case of the emergent categories.

#### **5.1 Framing Patterns in Television News Content**

##### ***5.1.1 Authoritative Voice Narration as a Tool of Legitimization***

The content analysis unveiled that there was uniformity in using a deep-toned, male voice-over narration, the pace of which was slow yet deliberate, and the diction was formal. This style was more dominant in reportages on national politics, economic developments, and diplomatic affairs. The narrators decided to use authoritative statements in more than 80 percent of the analyzed segments and often used idiomatic phrases like *The President underlined...*, *The state guarantees...* and *Kazakhstan remains the leader....* and this helped to convey the sense of dominance and power.

These types of voice framing assist the two functions of framing that Entman (1993), problem definition and treatment recommendation functions that remove narration not as informative but as a guiding force of how the audience should interpret a problem. Even the tone was set to provide a sense of trust in the authority of the institutions, and in line with that statement allegedly produced by Bucy and Newhagen (1999), that the vocal features like position, pacing, and pitch all make an impact on crowds assessments in terms of credibility.

### 5.1.2 Emotional Cueing Through Voice Modulation

The narration of sections related to the social issues was more emotional and touched on different tones; this might include medical areas, education, or human interest notes. In this case, narrators often reduced the tone of voice. They dropped the pace when sharing a tragic or bad experience. They utilized high pitch and more active rhythm during the triumphal ideas or communal occasions. This comes in opposition to the theory of mediated suffering advanced by Chouliaraki (2006), which adjures the use of vocals to draw emotional reactions in the audience.

In other instances, the report's narrators employed a nationalistic phrase such as "*our noble doctors*" or "*great Kazakh youth*," thus creating group identity through positive emotionalized support. According to the findings, although the tone of the voice does not directly imply a political message, it is still a type of persuasion, augmenting the emotional relevance of the narrative.

**Table 2: Frequency of key voice and visual framing techniques in analyzed news reports**

Framing Technique	Frequency (n=30)	Percentage (%)
Authoritative Voice Tone	24	80
Use of National Symbols	26	86.7
Emotional narration (soft tone)	18	60
Congruent Voice-Visual Framing	23	76.7

### 5.1.3 Symbolic Visuals and Repetitive National Imagery

In the news items sampled, visual framing was dominated by repetitive references to national symbols: flags, historical monuments, traditional clothing, and portraits of political leaders. These symbols were primarily used at the beginning of news reports and at the end, corresponding to framing bookends. This kind of iconography corresponds to that of Barthes (1977), who refers to such phenomena as the mythical image on the basis that recurring symbolic imagery is endowed with ideological significance and used to naturalize the state's power.

The structure of the images also included framing. The angles of the shot always put the government officials in a low position, making them look powerful. The simple citizens would be in a group shot, implying a sense of collectiveness and unity. Instead, protesters either were not depicted, or they were displayed in short, chaotic imagery scenes- often shot above or with fuzzy movement, which brings out disorder and instability. These are visual tactics that Coleman (2010) detected when analyzing media situations in politically reinforced control to manifest opposition as invalid or dangerous.



**Fig 1: Example of symbolic visual framing in a Kazakhstani news segment showing the president's speech backed by national symbols**

#### ***5.1.4 Synchronization of Voice and Visuals for Reinforcement***

The visual framing of the sampled news extracts showed repetitive national symbolism, with the recurring appearance of country symbols denoted as flags, monuments of the past, folklore costumes, and portraits of political leaders. These images were most commonly presented when making introductions and conclusions of a news report and served as framing bookends. This type of iconography follows Barthes's description (1977) of the mythical image or the repetition of symbolic images to the point that they gain ideological significance and have an effect of naturalizing state power.

The composition played a role in visual framing, too. Government officials were always filmed using low angles, which lent more authority to them, and citizens were usually shown in group shots, which implied unity and togetherness. In contrast, protesters were either unseen or depicted in short, violent clips that often appeared aerial or with in-focus movement in a frenzied style, depicting disorder and instability. These visual tactics are similar to those understood by Coleman (2010), who witnessed identical tactics used in the politically regulated forms of media to park opposition as unworthy or dangerous.

## **5.2 Audience Interpretations and Viewer Perceptions**

### ***5.2.1 Trust in Voice Narration Varies by Demographics***

The focus groups' responses showed that older people, especially reactors 45 and above years or older, would consider voice-over narration a possible source of believable and authoritative information. They linked the sound to professionalism and the reliability of states as the reflection of the norms of media consumption established in Soviet and early post-independence times (Kudaibergenova, 2016).

Conversely, the younger respondents were disbelieving about the commanding tone. Others called the narrators robotic or, one might say, scripted, with others seeing a lack of emotion. One of the participants said that trying to sell it to us a bit too hard. When the voice

is too good, I lose trust in it." The dissimilarity between the generations highlights this encoding/decoding model by Hall (1980), according to which various audiences interpret a standard message according to dissimilar cultural and ideological prisms.

**5.2.2 Emotional Framing Influence Viewer Sympathy**

All the participants noted that they were emotionally stimulated by the change in tone of voice, especially during human-interest stories. The narrator's soft, slow-paced speech made them sound more empathetic, thus enhancing trust and emotional connection. This confirms the claim by Geise and Baden (2014) that paralinguistic aspects of narration may direct moral judgment and emotional orientation.

Participants, however, raised some concerns that such techniques can be manipulative. One of the younger interviewees said, "They also slowly listened to the voice down and displayed a crying woman... as they want to make you feel guilty or proud, but it differs depending on the story."

**5.2.3 Visual Framing Shapes Perceptions of Power and Legitimacy**

Repeating national symbolism, ceremony shots, and flattering camera positions on the participants were identified by the participants as very general, but that was explained differently. The participants in rural locations and older viewers mostly interpreted these visuals as declarations of Kazakhstan's stability and improvement. In contrast, the urban and younger viewers considered them excessive and propagandist.

When presented with protest-connected footage, most respondents observed that images were blurred, short, or accompanied by a neutralizing voice-over. This gave the impression that such events were either in the margins or irrelevant. An Almaty University student said, "They showed the protest twice as long as they showed the smiling man of the government like nothing happened." This supports the argument by Messaris and Abraham (2001) that the omission of visuals and juxtaposition are effective as a framing device.

**5.2.4 Media Literacy and Language Shape Interpretation**

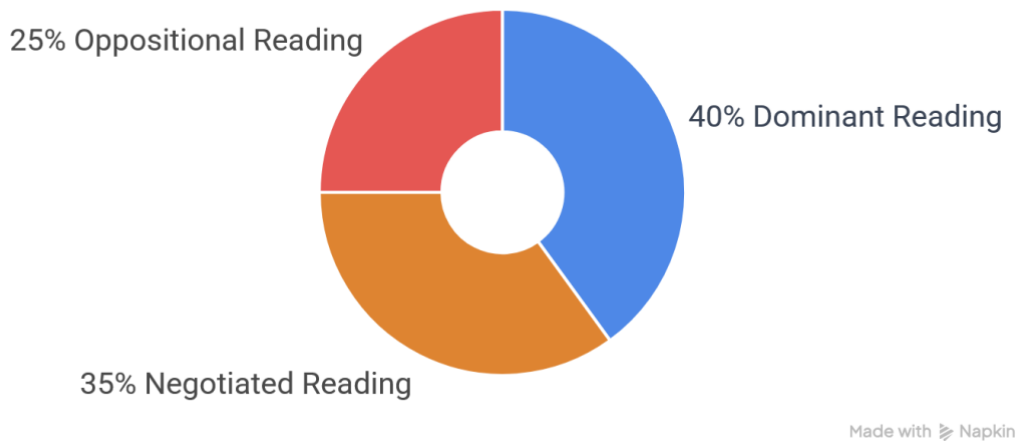
Lastly, frame interpretation was influenced by linguistic and educational background. The Russian-speaking audience was more prone to refute the narration, referring to international media and other sources of information. In the meantime, Kazakh-speaking respondents (and those in the countryside, in particular) were inclined to take the stories as being true and would frequently paraphrase the expressions as they are employed on air. This swerve connotes the idea of negotiated and oppositional reading, which Stuart Hall (1980) developed based on the audience's accessibility to competing discourses.

**Table 3: Interpretation of framed television news across audience subgroups**

<b>Demographic Group</b>	<b>Dominant Reading</b>	<b>Negotiated Reading</b>	<b>Oppositional Reading</b>
Older Rural Viewers	✓		
Younger Urban Viewers		✓	✓
Kazakh-Speaking Viewers	✓	✓	
Russian-Speaking Viewer		✓	✓

Figure 2 below illustrates how the focus group participants decoded the voice and visual frames. The dominant readings occurred more often in older and rural respondents, and the oppositional ones among younger, urban, and Russian-speaking subjects.

### Distribution of Audience Interpretation Categories in Kazakhstan



**Fig 2: Distribution of Audience Interpretation Categories in Kazakhstan**

### 5.3 Summary of Key Findings

The findings indicate that voice and image in Kazakhstani television news are well coordinated in narrating the contents in a manner that initiates state legitimacy, affective identification, and a sense of togetherness. Although these approaches are mostly successful with older and rural viewers, younger and urban audiences feel more skeptical, as they were influenced by media literacy and exposure to alternative stories.

These results indicate a rather complex interaction between media production strategies and consumer perception. There is nothing uniform in the effects of framing; they are context-specific and demographically varying.

## 6. DISCUSSION

The identified trends in this research give weighty empirical evidence to the idea that voice and visual framing in Kazakhstani television news are the key instruments in the construction of the perception of the viewer, particularly referring to the establishment of the correct relation of the subject in the area of institutional construct, national perception, and emotional state. The interrelationship of a heavy-handed narrative and symbolic images was not only a common occurrence, but very deliberate in that it created a logical story with ideological consistency. Notably, how highly the audience reacted to such framing strategies has been observed to be quite different across demographic aspects like age, language and media literacy count, thus reaffirming that media effects are not universal but influenced by social-cultural placement (Hall, 1980).

### 6.1 The Power of Authoritative Voice Framing

The fact that the narration all through the news articles is deep-toned and male oriented seems to agree with past observations that voice is both used in conveying information and

impression managing (Bucy & Newhagen, 1999). Within the process in Kazakhstan, where the governmental structures tend to pursue the idea of stability and legitimacy, the narrator's voice can be viewed as an extension of the state power. This paper establishes that these vocal devices, which appear confident, clear, infrequent, and formal vocabulary, are not a matter of style. Still, one of rhetoric used to convey competence and credibility.

However, according to the focus group discussions, there is a generation divide in the reception of these vocal strategies. The voice-over affects the older viewers by causing a sense of professionalism and truthfulness. In contrast, the younger viewers would rather see the voice-over as a performance or subject them to manipulation. This contrast reflects the paradigm theory of narrative offered by Fisher (1984), who identified audiences promulgating narratives according to their conclusions of coherency and correspondence to reality. Among a younger audience more media-literate the flawless, deadpan narration might not be a faithful portrayal of a story as it can seem false.

## 6.2 Emotional Manipulation and Affective Framing

The information also shows that voice-over narration is applied to hint at emotional reactions, be it softer intonation when stories of tragedy are to the fore or rapid pacing when narrating success stories of the nation. These activities are in line with Chouliaraki's (2006) model of the mediated effect, where audiovisual news frames are developed to elicit sympathy, pride, or fear.

Such emotional framing can increase engagement of viewers and readers although it also brings in ethical issues. Some viewers also disliked how narration and images were used to evoke certain feelings, specifically those related to nationalistic or paternalistic values. This implies a struggle between identification and detachment, which has some implications in the democratic discourse. When audiences feel manipulated by affecting their emotions, they might stage a withdrawal or foster a more defiant attitude towards the media, a process recorded in reception studies of state-controlled media systems (Votmer, 2013).

## 6.3 The Ideological Function of Visual Framing

The paper also concludes that the frequent use of symbolic images (flags, monuments, portraits, leader faces, and conventional clothing) establishes an image on which news stories are visually anchored in terms of national dignity and continuity. Such images are the kind that Barthes (1977) referred to as mythologies, that is, visible markers containing ideological messages purported to be normal within the cultural setting.

Such visual framing was the most evident during news reports presenting political speeches or state programs or reaction to crisis. As they continuously associate these images with positive texts, the broadcasters support their thinking about the world as state-centered where institutional actors are effective, well meant, and those easily accommodating. On the contrary, dissenting ideas or crises are usually depicted with the help of stock footage or are framed in the manner that does not reflect the scale of their importance.

The given practices are consistent with the functions of framing outlined by Entman (1993): a selection of reality to foster particular problem definition, causal explanation, and moral judgement. When applied to Kazakhstani news, the specifics of the role of framing would be its strict coordination of voice and image to create a harmonious, stable picture of the state, which in turn would support the existing power base.

#### **6.4 Audience Interpretation: Negotiated and Oppositional Readings**

The most intriguing feature in the audience data is the sense of heterogeneity in interpreting the frames. Here, we can directly apply the model of encoding/decoding developed by Stuart Hall (1980): the dominant reading may be taken by some viewers, especially those from rural backgrounds or older generations, whereas the negotiated reading or even the oppositional reading applies to the rest of the viewers.

Young, urban and Russian speaking respondents were inclined to criticize the coverage as too selective or propagandistic. They did not only doubt the narration but the visuals because they were judged to be either repetitive or unrealistic. This observation propagates the concept that media literacy and exposure to alternative information messages apply in the decoding process context (Livingstone, 2004). For example, foreign news consumers noticed loose ends or contradiction in the local news faster.

Interestingly, a lack of adjective consistency in dominant framing was also witnessed despite not always leading to the rejection of the news's contents. Many of them took a reflexive position, namely they acknowledged the strategic character of the framing but still involved themselves in the news to gain information or to live within a culture. This paradox can be explained by what Kraidy (2002) refers to as hybrid reception, which is the system of simultaneous internalization and rejection of messages in the media.

#### **6.5 Implications for Media Practices in Transitional Democracies**

The Kazakhstan media system is something of a hybrid between state control and commercial modernization models, which provides a unique background for researching the effect of framing. The rhetorical tactic of framing through voice and visual signs shows how media institutions manage the tension between keeping state discourses intact and appealing to more diverse and critical viewers. According to this research, it appears that. In contrast, framing techniques may effectively continue to work among dissimilar societies, they threaten to put off an increasing percentage of the populace which is less susceptible and globally entrenched. These findings are similar to Voltmer (2013) assertion that standardized framing strategy may lead to disintegrated reception effects in transitional democracies whereby institutions falter to meet the expectation of the new audiences. Furthermore, the persistence of authoritarian framing methods can restrict the media's attempt to engage citizens in an inclusive public debate. With change in the media systems, they should be more plural in their approaches to framing, where such approaches can be accommodative, encourage critical thinking, and not be emotionally blackmailed or ideological positioning.

### **7. CONCLUSION**

This paper explored the relationship between the narration of voices and the visual framing criteria in the television news in Kazakhstan, which depicted the significance of explaining how audiovisual elements use their meaning construction along with creating perception on the viewer of the semi-authoritarian media environment. Based on a mixed-method design, integrating content analysis of news broadcast and the focus group data, the study finds that narratives espoused by the state power, which corresponds to state legitimacy, national identity, and social cohesion, are actively prescribed with voice and image elements as the narratives are being presented by the TV news broadcast domain. The results illustrate that top-down male voice-overs and emotional tone of delivery are widely used to advise understanding and seek trust or compassion, respectively, depending on the subject. These narratives are further combined with the visual framing by the repetition of the symbols of the nation, camera angles that are favorable, and the limited portrayal of the opposition. More

importantly, the study proves the reason why synchronicity between narration and imagery is essential in enhancing persuasive effects, which is of theoretical magnitude as framing, narrative, and semiotic insights on the issue affirm.

Nevertheless, reaction to these methods is not unanimous in the audience. The reception of the intended framing across the older, rural, Kazakh-speaking viewers is very much acquiescent compared with younger, urban and Russian-speaking ones; the latter indulge in negotiated or counter readings. Such reception variety supports the coding/decoding explanation of Stuart Hall and the increased role of media literacy and availability of alternative discourses in media interpretation of news.

The present research is theoretically important to the studies of media on a global post because it aids in the expansion of the theories of framing and reception theories into a cross-cultural and post-Soviet context. In practice, it raises the moral implications of the producers of media in the transitional democracies. Emotional entrapment and the use of ideological imagery has been used by government, and is useful among some, but dangerous to other, more critical sectors of society. With the further development of the media in Kazakhstan, transparency, pluralism and the trust of the audience should become the leading principles of journalism in this country. The development of the effects of framing in political attitudes over a considerable period of time and over time must be the subject of future research on the subject. The means of applying these techniques is also to undergo changes as more digital and social media options become available. The comparisons of other similar transitional situations may also help to unveil more information on the universality and particularity of audiovisual framing strategies. At the end of this study, it is important to remark that in highly politically sensitive media systems, voice and visuals do not only serve as information carriers but as tools of power, that are deeply interlaced in the sociopolitical setting of communication as intended to publics.

#### REFERENCES

- Beissinger, M. R. (2013). Nationalism and the collapse of Soviet Communism. *Contemporary European History*, 18(3), 331–347.
- Bucy, E. P., & Newhagen, J. E. (1999). The micro- and macrodrama of politics on television: Effects of media format on candidate evaluations. *Journal of Broadcasting & Electronic Media*, 43(2), 193–210.
- Coleman, R. (2010). Framing the pictures in our heads: Exploring the framing and agenda-setting effects of visual images in political news. In P. D'Angelo & J. A. Kuypers (Eds.), *Doing news framing analysis* (pp. 233–261). Routledge.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58.
- Esser, F. (2008). Dimensions of political news cultures: Sound bite and image bite news in France, Germany, Great Britain, and the United States. *International Journal of Press/Politics*, 13(4), 401–428.
- Fisher, W. R. (1984). Narration as a human communication paradigm: The case of public moral argument. *Communication Monographs*, 51(1), 1–22.
- Freedom House. (2022). *Kazakhstan: Freedom in the World Report*. <https://freedomhouse.org>
- Geise, S., & Baden, C. (2014). Putting the image back into the frame: Modeling the linkage between visual communication and frame-processing theory. *Communication Theory*, 25(1), 46–69.
- Grabe, M. E., & Bucy, E. P. (2009). *Image bite politics: News and the visual framing of elections*. Oxford University Press.

- Hall, S. (Ed.). (1997). *Representation: Cultural representations and signifying practices* (Vol. 2). Sage.
- Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. University of Chicago Press.
- Junisbai, B. (2010). A tale of two Kazakhstans: Sources of political cleavage and conflict in the post-Soviet period. *Europe-Asia Studies*, 62(2), 235–269.
- Kim, Y. Y., & Barnett, G. A. (1996). The determinants of international news flow: A network analysis. *Communication Research*, 23(3), 323–352.
- Knobloch-Westerwick, S., & Hastall, M. R. (2006). Media framing and psychological needs: A study of individual differences and framing effects. *Communication Research*, 33(4), 481–506.
- Kudaibergenova, D. (2016). The ideology of development and legitimation of power in Kazakhstan. *Central Asian Survey*, 35(4), 421–437.
- Messariss, P., & Abraham, L. (2001). The role of images in framing news stories. In S. D. Reese, O. H. Gandy Jr., & A. E. Grant (Eds.), *Framing public life* (pp. 215–226). Lawrence Erlbaum Associates.
- Nurbekova, Z. (2018). Media independence in Kazakhstan: Between control and professionalism. *Central Asian Survey*, 37(3), 363–378.
- Open Society Foundations. (2021). *Mapping digital media: Kazakhstan*. Retrieved from <https://www.opensocietyfoundations.org>
- Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). SAGE.
- Schimpfössl, E., & Yablokov, I. (2020). *Propaganda and journalism in authoritarian states*. Routledge.
- Tashakkori, A., & Teddlie, C. (2010). *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. SAGE.
- Voltmer, K. (2013). *The media in transitional democracies*. Polity Press.