

Standard Language Ideology and the Construction of Northern Accent Prestige in Vietnamese Media

Nguyen Thuy Ngan^{1*}, Nguyen Minh Thu²

¹Diplomatic Academy of Vietnam

²UWC Dilijan College

DOI: <https://doi.org/10.51244/IJRSI.2026.1308000269>

Received: 05 September 2026; Accepted: 10 September 2026; Published: 21 September 2026

ABSTRACT

This study examined how Northern-accent prestige was constructed and naturalized in Vietnamese national television news and current-affairs programming. Rather than treating the elevated status of Northern speech as the self-evident result of linguistic clarity, the study approached standard pronunciation as an ideological and institutional achievement. It drew on critical discourse analysis, standard language ideology, and Bourdieu's concept of linguistic capital to examine a documented online pilot sample from four VTV1 programs: *Thời sự 19h*, *Chuyển động 24h*, *Việt Nam hôm nay*, and *Vấn đề hôm nay*. Comparison across program formats distinguished the concentration of standard-oriented speech in studio authority roles from the more heterogeneous voices heard in field reports and interviews. The analysis identified five mechanisms through which prestige was reproduced: institutional selection of the anchor voice; association of Northern-oriented pronunciation with seriousness, education, and national authority; framing of regional accents as local testimony rather than universal narration; correction and self-monitoring organized around clarity and professionalism; and repetition that converted a historically located variety into an apparently accent-less national norm. The analysis argues that Vietnamese television does not merely reflect a pre-existing hierarchy among accents. Through role allocation, scripting, editing, visual framing, and metalinguistic evaluation, it helped produce the Northern-oriented broadcast voice as the legitimate voice of public knowledge. Accent prestige was strongest where speakers-controlled topic definition, transition, interpretation, and closure, whereas regional diversity was most audible in short field-based testimony. The comparison further suggests that the distinction between a national voice and a regional voice depended not only on pronunciation but also on program genre, speaker role, and institutional positioning. The study therefore shifted attention from phonological difference to the discursive production of prestige and proposed a format-sensitive framework for examining accent hierarchy in institutional media. It concluded that a shared professional register could support national intelligibility without requiring regional conformity, provided that broadcasters separated communicative competence from inherited assumptions about the sound of authority.

Keywords: standard language ideology; Northern Vietnamese; accent prestige; media discourse; linguistic capital; language variation; television news; Vietnam

INTRODUCTION

Vietnamese displays extensive regional variation across Northern, Central, and Southern groupings and within provinces, cities, generations, and social groups. Differences involve consonants, vowels, tones, lexis, rhythm, and interactional style. Such variation is systematic, yet varieties circulate under unequal social conditions. In education, administration, language teaching, and national broadcasting, some forms are more readily associated with correctness, intelligence, authority, and professionalism. Educated Hanoi-oriented pronunciation has acquired a particularly strong claim to national legitimacy [1, 2]. Northern speech is often described as clearer, closer to orthography, or more widely understood. These properties cannot alone explain why one accent represents the nation. Meanings such as rationality, education, and trustworthiness emerge from history, institutional practice, and repeated representation. A voice repeatedly used to deliver

official information and interpret national events accumulates authority until its regional origin is heard as neutrality [3, 4]. National television is therefore a key site of standard language ideology. Broadcasting selects speakers, trains voice, allocates roles, edits recordings, and defines an acceptable public voice. The anchor is commonly heard as transparent and national, while an interviewee's pronunciation may index origin, authenticity, or informality. This contrast performs ideological work without explicit claims of superiority [5, 6].

This article examines how Vietnamese national television news constructs the prestige of Northern-oriented speech. It asks three questions. First, how is the broadcast voice organized in ways that make Northern-oriented pronunciation appear standard and institutionally authoritative? Second, what discursive mechanisms associate this voice with professionalism, neutrality, and national belonging? Third, how are regional voices included, marked, translated, or subordinated within national news discourse? The study did not rank accents or treat the North as homogeneous. Rural and provincial Northern varieties may also be stigmatized. Its object was a selective, educated, broadcast-compatible Northern orientation polished as standard Vietnamese. The analysis proposed five mechanisms: institutional selection, role allocation, indexical association, metalinguistic discipline, and repetition [7, 8].

LITERATURE REVIEW

Standard Language Ideology

Standard language ideology assumes that a language has, or should have, a single correct and superior form. Standardisation suppresses variability and presents the practices of socially powerful groups as neutral and universally intelligible. Standards are produced through schools, dictionaries, examinations, publishing, broadcasting, and judgments about appropriate speech [3, 4]. Standardization also works through erasure. The prestigious variety's history becomes invisible while other forms remain marked by geography, class, ethnicity, or informality. The standard speaker appears accent-less, allowing institutional preferences to be presented as technical requirements of clarity or professionalism [7, 9]. Irvine and Gal identify ionization, fractal cursivity, and erasure as recurrent processes. Linguistic features become signs of refinement or rusticity; national hierarchies recur within regions; and inconvenient diversity, including variation within Hanoi speech, is obscured [9].

Linguistic Capital and Media Authority

Bourdieu's linguistic capital explains why accents create unequal advantages. Schools, state agencies, and national media define legitimate usage, enabling speakers of authorized varieties to convert linguistic competence into credibility, employment, and symbolic authority. Others may invest additional labor in accommodation or concealment [10]. Broadcasting combines institutional selection with mass repetition. Presenters speak through the authority of the channel, studio, and news genre; visual centrality, scripted delivery, and technological quality reinforce their voices. Accent prestige is therefore simultaneously linguistic, visual, and institutional [10, 5]. Media participate in reinterment, through which recurring features become recognizable social styles. Repeated role assignment can associate accents with experts, leaders, villagers, or ordinary citizens; hierarchy arises from the roles attached to voices, not sound alone [11, 12].

Vietnamese Dialect Diversity and The Standard

Vietnamese is conventionally divided into Northern, Central, and Southern dialect regions, although variation is much finer. Phạm and McLeod document systematic phonological differences, while the ViMD dataset represents speech from all 34 provinces [1, 2]. Standard Vietnamese is nevertheless commonly associated with Northern speech and educated Hanoi usage. Northern cues are widely available through national media, and recent work describes Northern Vietnamese as the basis of the prestige variety. These studies document the hierarchy but do not fully explain its continual reproduction [13-15]. Northern and standard Vietnamese are not identical. Broadcast pronunciation may preserve distinctions reduced in everyday Hanoi speech. Here, Northern-accent prestige denotes an institutional orientation: pronunciation, lexical, prosodic, and delivery norms historically linked to the North and refined for national broadcasting

[1, 15]. Research on Vietnamese central television suggests that regional speakers may converge toward standard pronunciation on air. This supports a relational account: prestige belongs not to every Northern accent but to speech closest to an institutionally recognized model of educated national usage [8, 14].

Critical Discourse Analysis of The Broadcast Voice

Critical discourse analysis connects textual features with production practices and social power. For television news, it asks who may speak in which accent, in which role, and with what authority, not simply which accent occurs [5, 16]. The study therefore treats accent as part of discourse. Prestige is produced through scripting, camera position, turn-taking rights, and interpretive control, while regional voices become marked through location captions, short sound bites, and local testimony [5, 7].

METHODOLOGY

Research Design

The study adopts qualitative critical discourse analysis with a media-sociolinguistic focus. It examines how pronunciation interacts with speaker roles, program structure, visual framing, editing, and evaluations of professional speech [16]. This design was appropriate because standard language ideology rarely appeared as an explicit claim of superiority. It operated through routine choices: which voice introduced events, interpreted evidence, received extended time, or was described as clear and professional [3, 4].

Corpus Construction

The corpus consists of 12 publicly accessible VTVgo items from four programs: three each from Thời sự 19h, Chuyển động 24h, Việt Nam hôm nay, and Vấn đề hôm nay. VTV1 was selected because it functions as Vietnam Television’s national news and political-affairs channel. The purposive design retains the original comparison among a flagship bulletin, a studio-and-field news program, a news magazine, and an expert-led current-affairs format [17-21]. Inclusion required a stable VTVgo watch page, a displayed runtime, and sufficient archive information to identify the program and topic.

The pages were checked on 13 September 2026. Dates are recorded only when displayed in the title or expressly identified as a scheduled broadcast date; otherwise, the inventory reports “n.d.” The sample totals 189 displayed minutes. The unit of analysis for subsequent audiovisual coding is the continuous speaking segment. Appendix A provides the complete item-level inventory.

Table 1. Composition of the documented online pilot sample

Programme	Primary format	Items	Displayed minutes	Purpose in comparison
Thời sự 19h	Flagship evening bulletin	3	122	Studio authority and repetition
Chuyển động 24h	Studio-and-field social news	3	33	Field reporting and testimony
Việt Nam hôm nay	News magazine/social analysis	3	31	Regional and metalinguistic framing
Vấn đề hôm nay	Current-affairs previews	3	3	Policy and expert framing
Total	—	12	189	Documented online pilot

Source: VTVgo item pages listed in Appendix A. Runtime is the duration displayed on each page; access date: 13 September 2026.

Analytical Categories and Accent Identification

The coding framework contains five dimensions:

Accent is coded as a feature of the recorded performance rather than as a fixed personal identity. A Northern-oriented, Central-marked, or Southern-marked classification requires at least two recurrent phonological, tonal, lexical, or prosodic cues across sustained speech; isolated words, place of residence, personal identity, and geographic captions are insufficient. Speech that is too brief, noisy, mixed, or ambiguous is coded as indeterminate and excluded from accent-specific interpretation. These criteria prevent regional origin from being inferred solely from how a speaker is introduced.

(1) Speaker-role allocation distinguishes anchors, narrators, correspondents, officials, experts, and citizens to identify where interpretive authority is concentrated [5, 6].

(2) Markedness records whether location is highlighted through captions, introductions, or visual settings, while the studio voice is framed as national rather than regional [7].

(3) Discursive function distinguishes narration, explanation, evaluation, emotion, and testimony, asking whether regional voices interpret events or mainly authenticate local experience [5].

(4) Metalinguistic evaluation records references to clear, correct, natural, or professional speech and the normative assumptions embedded in them.

(5) Accommodation and editing identify moderation of regional features, style-shifting, and standard-oriented voice-over framing.

Table 2. Operational criteria for accent classification

Code	Decision rule	Permitted inference
Northern-oriented	At least two recurrent cues align with educated/broadcast Northern speech across sustained speech.	Performed Northern-oriented style; no inference about birthplace.
Central-marked	At least two recurrent phonological, tonal, lexical, or prosodic cues associated with Central varieties.	Audibly Central-marked in context; Central varieties are not treated as homogeneous.
Southern-marked	At least two recurrent phonological, tonal, lexical, or prosodic cues associated with Southern varieties.	Audibly Southern-marked in context; no inference about origin.
Mixed/accommodated	Cues from more than one orientation or systematic style movement within the segment.	Mixed realization or style-shifting in the recorded performance.
Indeterminate	Speech is too short, noisy, formulaic, inconsistent, or ambiguous.	Retained for role analysis but excluded from accent-specific claims.

Analytical Procedure and Ethics

Analysis proceeded in three stages: classification of speaker roles; close reading of introductions, transitions, quotation patterns, and studio-field relations; and interpretation through standard language ideology, linguistic capital, and indexicality. The corpus log records program, item identifier, date status, runtime, topic, source URL, regional framing, and discourse function. Only publicly broadcast material was used. Citizen interviewees were treated as participants in institutional media texts, not as objects of individual evaluation, and examples avoided reproducing accent mockery.

Direct observation was separated from interpretation: the broadcast or archive page establishes what was said, displayed, or institutionally framed, whereas explanations concerning recruitment, training, self-monitoring, audience attitudes, or editorial intention require independent evidence. The present revision does not report an inter-coder statistic because the pilot has not undergone independent second-coder checking. A follow-up coding round should have a second trained coder classify the complete pilot or a stated subset and document agreement and consensus decisions.

FINDINGS AND DISCUSSION

Program Format and The Distribution of Linguistic Authority

Observed program-format pattern. The four programs do not contribute identical kinds of speech. Thời sự 19h concentrates authority in a tightly scripted studio structure: anchors introduce national events, voice-over narration organises evidence, and field voices are inserted into a completed interpretive frame. Chuyên động 24h contains faster alternation between studio presentation, mobile reporting, and citizen testimony, producing greater audible regional diversity but retaining standard-oriented speech at points of introduction and synthesis.

Việt Nam hôm nay extends the news-magazine format, allowing presenters and narrators longer explanatory turns. Vấn đề hôm nay gives experts more speaking time, yet the host still controls topic definition, questioning, transitions, and final framing.

Prestige depended not only on frequency but on discourse position. A regional accent repeated in brief interviews did not carry the same authority as a studio voice that opened programs, explained causality, introduced statistics, and closed discussion. The analysis therefore distinguished exposure from authority. Program formats also shaped regional markedness. Studios appeared detached from locality, while field speakers were framed by place captions and local imagery. Expert titles could partly override locality, showing that prestige was jointly shaped by pronunciation, genre, role, and visual framing.

Table 3. Verified online evidence for the five proposed prestige mechanisms

Mechanism	Source	Verified evidence	Interpretation	Limit
Institutional selection	TS01-TS03	Thời sự 19h is described as VTV's most important daily bulletin, providing a comprehensive and official perspective.	The program voice is institutionally authorized.	The description does not establish presenter accent or recruitment practice.
Role allocation	VN02	The archive description defines the legal, health, and demographic consequences of early marriage in ethnic-minority areas.	Institutional discourse occupies a problem-defining role toward a geographically marked community.	Speaker turns, accents, and speaking time require video coding.
Indexical association	VD02	The economic-census preview emphasizes national scale, technology, method, data, and information that is “đúng, đủ, kịp thời.”	Policy programming is associated with expertise and epistemic authority.	A specifically Northern accent association is not established by page text.

Metalinguistic discipline	VN01	The description uses “phát âm không chuẩn,” “nói ‘ngọng,’” and “khắc phục lỗi L, N.”	Selected pronunciation is explicitly constructed as non-standard and correctable.	The archive description is not a speaker-level transcript.
Repetition	TS01-TS03	The same flagship bulletin identity appears on 31/7, 03/8, and 05/8/2026, with runtimes of 43, 39, and 40 minutes.	Institutional framing recurs across separately dated items.	Repeated phonetic features require auditory coding.

Institutional Selection: The Standard Voice as an Occupational Requirement

Observed institutional pattern. The first mechanism constructing Northern-accent prestige is institutional selection. Across the four formats, the most institutionally controlled positions are the studio anchor in Thời sự 19h, the principal presenter in Chuyển động 24h and Việt Nam hôm nay, and the host who frames expert discussion in Vấn đề hôm nay. These speakers open and close segments, announce transitions, establish topic significance, and determine how field material is interpreted. Selection for these roles therefore establishes more than an aesthetic preference: it identifies a voice as capable of representing public truth. The expanded corpus makes this hierarchy visible across formats rather than within a single bulletin.

Interpretation. The prestigious broadcast voice is not identical to spontaneous Hanoi conversation. It is carefully paced, lexically controlled, and prosodically disciplined. Sentence boundaries are clearly projected; emphasis follows informational structure; emotional display is restrained; and potentially ambiguous words receive deliberate articulation. These characteristics are genuine professional skills. The ideological issue arises when a trained performance is interpreted as the natural property of a regional accent. Professional scripting, rehearsal, audio technology, and institutional authority disappear from perception, leaving the impression that the accent itself is inherently clear and serious.

Institutional explanation requiring independent evidence. Recruitment and training reinforce this association. Aspiring presenters learn that successful broadcasting requires a “clear,” “round,” or “standard” voice. Such descriptors appear neutral because every broadcaster needs intelligibility. Yet the operational meaning of clarity is rarely accent-free. It is calibrated against an existing model produced by earlier generations of prestigious presenters. Applicants whose speech already approximates the model require less accommodation and are more easily heard as naturally professional. Others must demonstrate that they can reduce features treated as local [5, 6].

Interpretation. This process resembles Bourdieu’s linguistic market. The institution rewards speakers who possess, or can acquire, the authorized pronunciation. Their accent becomes linguistic capital because it grants access to visible professional roles. The resulting overrepresentation then confirms the belief that the accent is best suited to those roles. Selection and perception sustain each other in a circular process [7]. No explicit policy excluding non-Northern speakers was necessary. Inherited ideas of clarity and correctness could guide apparently technical judgments, so the question of whether a candidate could deliver national news clearly also implied whether the candidate resembled the established national voice [5].

Role Allocation: Narrators of the Nation and Witnesses from Regions

Observed role-allocation pattern. A second mechanism is unequal role allocation. In a typical national news sequence, the studio presenter introduces the topic and defines its significance. A field report then provides images, narration, and brief quotations. Local residents may describe effects on their lives; provincial officials may provide figures; experts may interpret causes; and the reporter closes with a summary. Although multiple accents can be heard, they do not necessarily carry equal discourse authority. The anchor’s speech is de-localized. The presenter is physically located in a studio, but that location is not framed as a region competing with others. The studio symbolizes the centre from which the nation becomes visible. By contrast, field voices are introduced through place names and visual signs: a province caption, a

landscape, local clothing, a flooded road, a factory, or a market. Their accents are therefore heard together with geographic specificity [11, 12].

Interpretation. This structure produces an asymmetry between “voice from nowhere” and “voice from somewhere.” Northern-oriented studio speech can operate as the transparent medium of information, while Central or Southern speech often indexes a locality. The distinction is not absolute. Many reporters and officials with non-Northern backgrounds speak in national media, and regional correspondents may perform extended analytical roles. Nevertheless, even inclusion can preserve hierarchy when local voices supply experience and the central voice supplies interpretation [18, 3].

Illustrative analytical sequence. Consider disaster reporting. A local resident may speak emotionally about damage in a marked regional accent. The testimony provides authenticity and human impact. Immediately before and after the quotation, however, a standard-oriented voice-over explains the scale of the event, connects it to governmental action, and guides the audience’s interpretation. The local voice is indispensable but bounded. It functions as evidence within a narrative controlled elsewhere [10]. Economic reporting showed a similar division: workers and consumers offered short evaluations, while presenters introduced trends and policy significance. Accent hierarchy therefore coincided with epistemic hierarchy: some voices interpreted the nation, while others experienced a part of it.

Indexical Association: Seriousness, Education, and Neutrality

Interpretation. The third mechanism is indexical association. Through repeated co-occurrence, Northern-oriented broadcast speech becomes linked with seriousness, education, bureaucratic competence, and political authority. These meanings are not encoded in phonology. They accumulate because the accent is repeatedly heard in settings already invested with prestige: news studios, official ceremonies, public announcements, expert interviews, and educational institutions [3, 4]. The association becomes especially strong in hard-news genres. A controlled Northern-oriented voice commonly introduces government policy, diplomatic activity, financial indicators, public security, and national commemorations. The formal visual field: desk, screens, maps, graphics, official titles support the voice. Over time, viewers may hear the accent as inherently appropriate for consequential information [10].

Interpretation. Neutrality is central to this process. News organizations value a voice that does not distract from content. Yet “non-distracting” is socially learned. An accent becomes unobtrusive because audiences have repeatedly encountered it in authoritative roles. A less institutionally familiar accent may attract attention not because it is objectively less intelligible but because listeners have learned to hear it as marked [11]. This helps explain why the standard is often described as universally understandable. Comprehensibility matters, and phonological differences can affect perception. However, comprehension is also shaped by exposure. National audiences encounter Northern-oriented speech extensively through schooling and media. Familiarity is then reinterpreted as intrinsic clarity. The media variety benefits from the very exposure that its alleged clarity is used to justify [11]. The indexical field was also classed and gendered. Broadcast voices co-occurred with educated vocabulary, managed appearance, and urban professional identity. Prestige thus indexed access to training and institutional recognition, not simply Northern geography.

Metalinguistic Discipline: “Clarity” As an Ideological Keyword

Metalinguistic analysis. A fourth mechanism operates through metalinguistic discourse: the language used to evaluate language. Terms such as “clear pronunciation,” “standard voice,” “correct articulation,” and “professional delivery” establish a moral and technical vocabulary for regulating speech. They allow institutions and speakers to discuss accent without appearing to rank regions [13]. Clarity is necessary in broadcasting, but it is not a self-explanatory criterion. A speaker can be comprehensible while retaining regional phonological features. When clarity is equated with movement toward one model, the concept becomes ideological. It shifts from a communicative outcome, whether audiences understand to a conformity requirement, whether the voice resembles the authorized form.

Institutional explanation requiring interview or reception evidence. The language of correction also changes how speakers understand themselves. A Central or Southern reporter may describe accommodation as “fixing” the voice rather than learning an additional professional style. The distinction matters. Style expansion treats speakers as adding a resource while retaining legitimate identities. Correction implies that the original accent contains defects. This discipline can produce self-monitoring. Speakers anticipate judgments and suppress salient features before any editor intervenes. In Bourdieu’s terms, the linguistic market is internalized. The institution’s standards operate through the speaker’s own bodily control: tongue placement, tonal contour, tempo, and lexical choice. Successful accommodation may create professional opportunity, but it can also make the labor of non-standard speakers invisible [10]. Audience criticism further reinforced the norm. Digital comments on pronunciation and regionalisms made viewers an informal regulatory body: slips by prestigious presenters appeared exceptional, whereas regional features could be treated as evidence of unsuitability [3].

Repetition and Naturalization

Interpretation. The fifth mechanism is repetition. National television news is serial and habitual. The same programme titles, opening music, studio layouts, turn-taking patterns, and vocal styles recur daily. Repetition converts institutional choice into auditory common sense [22]. A viewer does not need to endorse an explicit doctrine of Northern superiority. Daily exposure teaches a subtler association: this is what national information sounds like. Children encounter the voice before they possess sociolinguistic terminology. Adults hear it during crises, elections abroad, diplomatic visits, storms, market changes, and public celebrations. The accent becomes attached to the temporal rhythm of national life. This process is naturalization. A historically contingent broadcast norm appears inevitable because alternatives are rarely heard in equivalent roles. When regional voices do appear, they are often presented as welcome diversity within a framework whose center remains unchanged. The standard is therefore reproduced not only by excluding difference but by organizing difference around itself [7]. Repetition also creates perceptual accommodation. Listeners from all regions become skilled at understanding Northern-oriented speech. This is a practical national resource. However, accommodation is uneven: speakers of less represented varieties must often understand the standard and modify their own speech, while standard-oriented speakers face less pressure to understand or use regional forms. National intelligibility is achieved through asymmetrical linguistic labour [18, 3].

Controlled Pluralization in Contemporary Broadcasting

Counter-pattern. Vietnamese national television is not linguistically uniform. Regional correspondents, experts, officials, athletes, artists, and citizens bring diverse voices to the screen. Some programs deliberately foreground cultural diversity, and contemporary audiences may value presenters who sound more conversational and less ceremonial. Digital competition also encourages broadcasters to adopt flexible styles. These developments complicate any simple claim that non-Northern accents are absent. The more important question is whether they appear with equal authority. Controlled pluralization occurs when diverse voices are included but assigned differentiated functions. A regional accent may be accepted in live reporting, entertainment, weather, or local-interest stories while the central hard-news anchor remains closer to the inherited standard. Diversity is visible, yet the prestige structure persists [11].

These boundaries could shift when regional correspondents repeatedly performed authoritative roles. Online media increased exposure to diverse voices, but speech technologies could reproduce hierarchy when trained disproportionately on standard speech. Dialect-inclusive datasets remained necessary to prevent non-standard voices from being treated as noisy inputs [2].

From Phonological Difference to Discursive Power

Synthesis. The findings suggest that Northern-accent prestige cannot be explained simply by counting phonological distinctions. Northern speech may preserve contrasts useful for national intelligibility, but structural properties acquire social significance only through institutions. A pronunciation becomes prestigious because it is selected, amplified, and attached to authoritative roles [11, 4]. The analytical model

can be summarized as a cycle. Historical and political centrality shapes a broadcast norm. Recruitment and training privilege speakers who approximate it. Those speakers occupy high-authority roles. Repeated exposure associates their voices with expertise and neutrality. Audience expectations then feed back into recruitment and self-monitoring. The result is a self-reinforcing linguistic market [11, 7].

Standardization could therefore be useful and unequal. A shared register facilitated communication, but inequality arose when it became a hierarchy of persons and regions. A pluralistic model would distinguish intelligibility from conformity and assign regional voices interpretive and anchor roles, not only testimonial ones [11].

IMPLICATIONS

The findings have implications for broadcasting, professional training, education, and language technology.

Broadcasters should evaluate accent diversity through role distribution, not presence alone: who anchors hard news, narrates voice-overs, appears as an expert, or remains confined to local testimony.

Presenter training should separate accuracy, pacing, preparation, and intelligibility from regional conformity. Speakers could develop a formal broadcast register without treating home accents as errors [3, 11].

Education should present the national media voice as a shared formal variety rather than the only correct Vietnamese, reducing linguistic insecurity while preserving a common professional register [1].

Speech technologies required dialect-inclusive corpora and transparent error testing; otherwise, systems trained mainly on standard voices risked converting symbolic prestige into infrastructural advantage [2].

Future research should combine matched-guise experiments, interviews with journalists, and longitudinal program analysis to test audience evaluations, accommodation practices, and changes in accent diversity.

CONCLUSION

The analysis suggests that the prestige of Northern-oriented speech in Vietnamese national television is institutionally and discursively constructed. Broadcasting selects particular voices, associates them with authoritative roles, evaluates them through clarity and professionalism, and repeats them until a regional-historical variety appears neutral and national. Five mechanisms sustain this process: institutional selection, unequal role allocation, indexical association, metalinguistic discipline, and repetition. They distinguish the apparently accentless narrator of the nation from geographically marked regional speakers. Within the proposed framework, regional voices often authenticate experience, while the studio voice retains interpretive control. A shared broadcast register remains practically valuable, but professional standards should separate communicative effectiveness from regional hierarchy. Multiple accents can convey expertise and national authority as well as local authenticity. The study therefore shifts attention from ranking phonological systems to the institutional positions from which voices are heard. Vietnamese media may therefore do more than transmit a standard; through selection, framing, and repetition, they continually produce it.

Appendix A. Online Corpus Inventory

All pages were accessed on 13 September 2026. “n.d.” indicates that the page did not display a reliable broadcast or publication date. A contextual date in a description is not treated as an upload date.

ID	Programme	Date/status	VTVgo URL	Runtime	Topic/frame
TS01	Thời sự 19h	31/7/2026	https://vtvgo.vn/watch/52n0	43 min	Daily national and

		(title)			international news
TS02	Thời sự 19h	03/8/2026 (title)	https://vtvgo.vn/watch/53Hi	39 min	Daily national and international news
TS03	Thời sự 19h	05/8/2026 (title)	https://vtvgo.vn/watch/513N	40 min	Daily national and international news
CD01	Chuyển động 24h	n.d.	https://vtvgo.vn/watch/4hYy	3 min	Flood impacts and mutual aid
CD02	Chuyển động 24h	01/10/2025 (context)	https://vtvgo.vn/watch/4kl0	10 min	Road-toll account conversion
CD03	Chuyển động 24h	n.d.	https://vtvgo.vn/watch/4i66	20 min	National reunification commemoration
VN01	Việt Nam hôm nay	n.d.	https://vtvgo.vn/watch/4hLs	13 min	Correction of L/N pronunciation
VN02	Việt Nam hôm nay	n.d.	https://vtvgo.vn/watch/4lbL	9 min	Early marriage in ethnic-minority areas
VN03	Việt Nam hôm nay	October 2025 (context)	https://vtvgo.vn/watch/4pA0	9 min	Flood recovery in Central and Northern regions
VD01	Vấn đề hôm nay	04/8/2025 (scheduled)	https://vtvgo.vn/watch/4kSz	1 min	Local disaster-response capacity
VD02	Vấn đề hôm nay	07/01/2026 (scheduled)	https://vtvgo.vn/watch/4q8D	1 min	Economic census and data policy
VD03	Vấn đề hôm nay	11/8/2025 (scheduled)	https://vtvgo.vn/watch/4kXz	1 min	Policy for attracting international talent

Appendix B. Evidence-Status Note

Table 3 reproduces wording available on official VTVgo pages. It provides concrete evidence of institutional framing, geographic marking, language evaluation, and cross-page repetition. It does not replace broadcast transcripts or auditory accent coding. Accordingly, speaker counts, speaking time, timecodes, and phonetic classifications are not reported as if they had been observed. These fields should be added only after the recordings have been manually segmented and independently checked.

Author's Biography

Nguyen Thuy Ngan is a lecturer in the Faculty of English at the Diplomatic Academy of Vietnam and a doctoral researcher in Political Science. Her research interests include sociolinguistics, language ideology, media discourse, public policy, demographic ageing, and institutional change. Her current work examines how language, power, identity, and institutional authority interact in Vietnamese public communication, alongside comparative research on sustainability and social policy.

Nguyen Minh Thu is a student in the International Baccalaureate Diploma Programme at UWC Dilijan College. Her academic focus centers on the dynamic interplay between community formation, cultural identity, and shared linguistic practices.

REFERENCES

1. Pham B, McLeod S. Consonants, vowels, and tones across Vietnamese dialects. *Int J Speech Lang Pathol.* 2016;18(2):122-134. doi:10.3109/17549507.2015.1101162.
2. Nguyen VD, Dang TC, Nguyen LT, Nguyen KV. Multi-dialect Vietnamese: task, dataset, baseline models, and challenges. In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing.* Association for Computational Linguistics; 2024. p. 7444-7458.
3. Lippi-Green R. *English with an Accent: Language, Ideology, and Discrimination in the United States.* 2nd ed. New York: Routledge; 2012.
4. Milroy J, Milroy L. *Authority in Language: Investigating Standard English.* 3rd ed. London: Routledge; 1999.
5. Fairclough N. *Media Discourse.* London: Edward Arnold; 1995.
6. van Dijk TA. *News as Discourse.* Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
7. Gal S, Irvine JT. *Signs of Difference: Language and Ideology in Social Life.* Cambridge: Cambridge University Press; 2019.
8. Nguyen VL. Local accents on Central Television of the Socialist Republic of Vietnam. *Russ J Vietnamese Stud.* 2018;2(4):66-78. doi:10.24411/2618-9453-2018-10038.
9. Irvine JT, Gal S. Language ideology and linguistic differentiation. In: Kroskrity PV, editor. *Regimes of Language: Ideologies, Politics, and Identities.* Santa Fe, NM: School of American Research Press; 2000. p. 35-84.
10. Bourdieu P. *Language and Symbolic Power.* Cambridge, MA: Harvard University Press; 1991.
11. Coupland N. *Style: Language Variation and Identity.* Cambridge: Cambridge University Press; 2007.
12. Johnstone B, Andrus J, Danielson AE. Mobility, indexicality, and the enregisterment of Pittsburghese. *J Engl Linguist.* 2006;34(2):77-104. doi:10.1177/0075424206290692.
13. Brunelle M. Tone perception in Northern and Southern Vietnamese. *J Phon.* 2009;37(1):79-96. doi:10.1016/j.wocn.2008.09.003.
14. Le T, Tuan LA. A parallel corpus for Vietnamese Central-Northern dialect text transfer. In: *Findings of the Association for Computational Linguistics: EMNLP 2023.* Association for Computational Linguistics; 2023. p. 13839-13855.
15. Tran TT. Language ideologies in Vietnam. In: Shei C, Li S, editors. *The Routledge Handbook of Asian Linguistics.* London: Routledge; 2022. p. 495-514. doi:10.4324/9781003090205-34.
16. Fairclough N. *Critical Discourse Analysis: The Critical Study of Language.* 2nd ed. London: Routledge; 2013.
17. Vietnam Television. VTV1 news and current-affairs programming [television programs and digital archive]. Accessed July 2026.
18. Vietnam Television. Chuyển động 24h chính thức thay đổi khung giờ phát sóng từ 1/1/2026. VTV.vn. 2026.
19. Vietnam Television. Chuyển động 24h: program archive. VTV.vn. 2026.
20. Vietnam Television. Việt Nam hôm nay: program archive. VTV.vn. 2026.
21. Vietnam Television. Vấn đề hôm nay: program archive. VTV.vn. 2026.
22. Silverstein M. Monoglot standard in America: standardization and metaphors of linguistic hegemony. In: Brenneis D, Macaulay RKS, editors. *The Matrix of Language: Contemporary Linguistic Anthropology.* Boulder, CO: Westview Press; 1996. p. 284-306.