

Multimodal Manipulation In Film Dialogue

Tulanbaeva Shahnozahon Shavkatbekovna

Lecturer at Andijan State Institute of Foreign Languages, Uzbekistan

Received: 31 July 2025; **Accepted:** 28 August 2025; **Published:** 30 September 2025

Abstract: A discourse-pragmatic analysis was conducted on film dialogue manipulation, integrating semantic-pragmatic and multimodal perspectives. Verbal tactics (euphemism, framing labels), prosodic cues (intonation, timing), bodily signals (gesture, gaze), and figurative tropes (metaphor, hyperbole) were annotated in four English-language films involving power asymmetries. Manipulative influence was found to arise from coordinated device clusters rather than single markers. In particular, lexical reframing was anchored by spatial-kinetic cues, and persuasive prosody was used to make assertions “non-negotiable.” The analysis yielded a comparative device-function typology table and an analytical flowchart, contributing tools for studying discourse-based manipulation in multimedia contexts.

Keywords: Film discourse; manipulation; pragmatics; multimodality; prosody; gesture; metaphor; euphemism; audience design.

Introduction: Manipulation was approached as a discourse phenomenon in which interpretive options are constrained for the recipient while the speaker maintains a facade of reasonableness. Unlike straightforward persuasion, manipulation often relies on power asymmetries and implicit control over context. In discourse analysis, manipulation has been linked to strategic framing, information control, and ideological reproduction. This study viewed film dialogue as a laboratory for such strategies, since cinematic speech is crafted and edited yet presented as spontaneous conversation.

Film discourse is inherently multimodal: dialogue combines with visuals and performance. Viewers occupy a double role—they are addressed both as characters in the story and as a real-world audience. Kozloff (cited in Dynel) observed that film spectators are “no longer passive eavesdroppers; they are ratified listeners,” privy to characters’ unspoken intentions. This recipient-oriented design means that manipulative dialogue performs two tasks: intra-diegetic influence (character-to-character) and extra-diegetic audience alignment.

The present research we questioned following: What semiotic resources are used to achieve manipulation in film dialogue, and how are these devices functionally

grouped? A mixed-method, discourse-analytic procedure was followed. First, relevant scenes were selected from four popular films involving power/play. These scenes were transcribed (from subtitles and available scripts) and segmented into “moves.” Each move was annotated for four device categories: lexical-stylistic (e.g. euphemism, naming), phonational (intonation, tempo, voice quality), kinetic (gesture, gaze, proxemics, object reference), and tropological (metaphor, hyperbole, irony). These groups were then linked to manipulative functions such as framing realities, legitimating power, coercion, and emotional appeal. The analysis was grounded in prior work on discourse and power, pragmatics, and multimodality.

Subsequently, device functions were compared across films to identify patterns. The outcome was an analytical flowchart (presented below) and two tables: one typologizing devices and functions, and one cataloguing film examples (with approximate timecodes and sources). The coding was performed blind to viewer judgments, focusing on systematic criteria (e.g. recognized euphemisms), authoritative intonation patterns. Timecodes are given as approximate (edition-dependent) locators.

METHOD

Discourse and manipulation

Teun van Dijk defined discourse-level manipulation as “social power abuse” that interferes with recipients’ reasoning [19, p. 360]. It operates by controlling topicality, presuppositions, and stance while appearing rational [19, pp. 359–363]. Accordingly, this study treated manipulation as an interactional constraint: speakers use linguistic and para-linguistic means to steer how listeners construct reality. Key implications for analysis were: (a) look beyond literal truth/falsity to framing, omission, and emphasis [19]; (b) consider the institutional context and role-relationships that license certain moves [5; 19]. This perspective is consistent with critical discourse analysis, where language choice is viewed as ideological action.

Speech-act theory and politeness principles are relevant but not fully determinative. We adopt Searle’s distinction that assertives and directives underlie what is treated as “the case” vs. “to be done” [14]. In manipulation, assertives might be half-asserted or masked, and directives may be hidden as suggestions. Brown and Levinson’s politeness theory [4] implies that face-threatening intentions (e.g. orders) can be packaged as “rational” requests or euphemisms. Thus manipulation often exploits polite forms or wit to avoid overt force while still pressuring compliance [4].

Film dialogue differs from natural conversation. It is scripted, condensed, and performed. Film dialogue has been described as an “illusion” of ordinary talk, where every pause and word is designed [10, 71]. Alena Kačmárová notes: “A film script and its portrayal through acting, technical components and the viewers’ constant reference to the dramatic context create the illusion of natural/idiomatic linguistic and non-linguistic behaviour” [10, 71]. In practice, this means the analyst has access to the entire context (editors’ choices, camera focus) rather than just the text. Hence manipulation in film can be meticulously documented: we know when and how scenes were staged to influence both characters and spectators.

Moreover, from a multimodal viewpoint, dialogue in film must be seen as jointly produced with gesture, facial expression, camera framing, and sound. For example, an authoritative claim (“Because I said so”) may be undercut or reinforced by a speaker’s smug smile, a pointed look, or ominous background music. Martyn Ford’s concept of “audience design” applies: characters are implicitly performing for the viewer as much as addressing each other [6,1629]. Thus, the viewer’s comprehension is guided by what the narrative shows (e.g. a character’s nervous glance) and says.

Contemporary discourse analysis recognizes that meaning is often distributed across modes. A systemic-

functional approach (O’Halloran) sees visual imagery and speech co-constructing ideational meaning [16,443–447]. Kress and van Leeuwen’s visual grammar similarly highlights composition and visual framing as semantic [13]. Hence we treated manipulative meaning as emerging from orchestrated cues rather than a single channel. For instance, emphasis might be signaled by raised voice and a sharp hand gesture simultaneously; both must be noted to understand the pragmatic intent.

In particular, two non-verbal strata were foregrounded:

Prosody (phonation): Variations in intonation, loudness, rate, and pausing were annotated. These are known to index emotional intensity and speaker stance [9,1]. For example, a lowered pitch and slow pace can convey menace, whereas a clipped tone may convey command. These cues were noted from context (film audio) and confirmed against scripts when available.

Kinesics (gesture, gaze, posture): Following Kendon’s work [11,11], gestures were treated as “visible action as utterance.” All apparent deictic gestures, emblematic gestures (e.g. chopping motion), and eye contact (or avoidance) were recorded, as they function as interactive acts. If a character points at another while speaking, that gesture was coded alongside the spoken turn. Spatial relations (e.g. looming close or turning away) were also considered influential.

Device typology

To structure the analysis, we adopted four broad device categories (some overlap is inevitable, but they served as analytic tools):

Lexical-stylistic devices: Choice of words and expressions including euphemism vs. dysphemism [5,6], loaded labels, jargon, hyperbole. These perform framing (labeling issues as “necessary” vs. “optional”) and set the ideological tone. Example: calling a fee an “administrative cost” (euphemism) or an “insurance scam” (dysphemism). This category captures tactics often studied in discourse manipulation.

Phonational (prosodic) devices: Intonation patterns, tempo changes, loudness, and strategic pausing. These tune the force of an utterance: e.g. a final falling tone often signals finality or seriousness. Prosody was coded following intonational pragmatics, and known cues (e.g. “authoritative voice”).

Kinetic (embodied) devices: Body language including gesture (pointing, beckoning, illustrative motion), facial expression, and proxemic behavior. Also included are interactions with objects in the scene (handing a document, stopping someone physically). Kinetic cues act as non-verbal signals of dominance or solidarity and

can emphasize or contradict speech.

Tropes (figurative patterns): Metaphor, metaphorical imagery, hyperbole, irony, and narrative analogies. Following Lakoff and Johnson [14], metaphors (e.g. framing politics as “war”) were considered as ways to map one domain onto another, simplifying complex persuasion. Tropes were noted both in verbal and visual form (e.g. an onscreen visual metaphor, or a spoken hyperbole).

These groupings provided the axes for comparison. In the analysis below, we often refer to “a lexical-framing device” or “a tropological move” to indicate which category is at play.

Four English-language films were chosen because they centrally dramatize manipulative communication in social institutions:

“Wag the Dog” (dir. Levinson, 1997) [15] – political spin and media manipulation.

“Thank You for Smoking” (dir. Reitman, 2005) [17] – corporate public relations and rhetorical sleight-of-hand.

“Devil Wears Prada” (dir. Frankel, 2006) [8] – fashion publishing hierarchy and workplace coercion.

“The Wolf of Wall Street” (dir. Scorsese, 2013) [18] – sales pressure and charisma in financial markets.

From each film, scenes were selected using these criteria: (i) a clear power imbalance or persuasive goal, (ii) significant dialogue exchange (more than one turn), (iii) visibility of gestures or props in the scene. We focused on professional or political contexts (thus excluding purely domestic scenes).

Subtitles from official releases (English SRT files) were used to transcribe the dialogue turns, with manual correction as needed for idiomatic phrasing. When available, published scripts (e.g. “Devil Wears Prada” by Aline Brosh McKenna [20]) were consulted for exact wording and any stage directions.

RESULTS AND DISCUSSIONS

Each selected scene was broken into speaker turns and, when large turns contained multiple acts, into smaller “moves” (e.g. introduction of a topic, a clarifying question). Each move was then annotated along the four device dimensions:

Lexical: Identification of any notable phrasing. For example, euphemisms (“restructuring” for layoffs) or dysphemisms (“junk” for ineffective product) were flagged. Idiomatic phrases that might manipulate perception (e.g. “selling us short”) were noted. Each marked phrase was assigned a rough type (metaphor,

label, euphemism, etc.) and function (e.g. “legitimizing downsize as efficiency”).

Phonational: Based on script and listening, each utterance was characterized by its delivery: we looked for key stressed words, pitch jumps, pauses at clause boundaries, and speech rate. For instance, a speaker may slow down on a crucial noun to add weight, or raise pitch at end to signal a challenge. These were described qualitatively (e.g. “low, measured tone, pausing after each clause”).

Kinetic: Using video playback (when possible) and script descriptions, visible actions were logged. This included gestures that accompanied speech (pointing, hand chopping), gaze shifts (e.g. eye contact maintained or broken), posture changes (leaning in), and movements (stepping closer, turning away). We also logged conspicuous uses of props (handing an item, showing a document) or environmental cues (lights dimming, a character's picture displayed). Each kinetic cue was paired with the co-occurring speech to assess its pragmatic role (e.g. “pointing at contract while saying “sign here” as directive intensifier”).

Tropes: Whenever metaphorical or ironic language was used, it was noted. For example, battle metaphors (“we’re under attack”) or exaggerations (“I’d sooner die”) were coded as such. We also noted if a speaker used narrative analogies (“this is our Woodstock”). These were cross-referenced with existing lists of metaphors in propaganda [14].

After annotation, each device instance was linked to one or more functions derived from the manipulation literature. These functions included: framing (setting the interpretive perspective), legitimation (justifying an action or authority), coercion (applying pressure or threat), and aesthetic/affective appeal (engaging emotion, humour, distraction). For example, describing a firing as “rightsizing” (lexical euphemism) was coded as “framing the downsizing as positive necessity.” Lowering voice when giving instructions was coded as “coercive register.”

The coding was performed iteratively by at least two analysts (the author and a research assistant) to ensure consistency. Disagreements were resolved by discussion.

The following table organizes the four device categories against common manipulative functions, with illustrative film-based examples (brief quotes or descriptions).

Table 1. Categories of manipulative devices, their discourse functions, and film discourse examples. (Bracketed sources document theoretical claims.)

Device group	Description	Typical Manipulative Functions	Film example (short)	Key source
Lexical–Stylistic	Word choice and phrasing (euphemism, dysphemism, labels, jargon, loaded terms) [5, p. 6].	Framing: redefining concepts (e.g. “collateral damage” vs “victims”); Legitimation: normalizing action (e.g. “downsize” vs “fire”); Stigmatization: negative labeling of dissenters.	“It’s actually cerulean, [not turquoise]” – normalizing authority of fashion terms [8]	[5]; [19]
Prosodic/Voice	Intonation contour, speech rate, loudness, pauses, voice quality (drawn from Brazil [3] and Guellai et al. [9]).	Authority display: measured, low-tone speech; Urgency/Threat: abrupt, loud tone; Rapport-building: soft, even pitch; Turn-control: strategic pausing to claim floor.	Miranda Priestly’s clipped tone in command – “Details of your incompetence do not interest me.” [8]	[3]; [9]
Kinetic (Gesture)	Body movement, gesture, gaze, proxemics, object manipulation [11]. Includes handling of props as communication.	Dominance/Intimidation: looming posture, direct gaze; Reinforcement of speech: beat gestures on keywords; Visual metaphor: using objects (e.g. pen as “cash symbol”); Compliance pressure: pointing or moving into listener’s space.	Pointing at the audience while saying “Sell me this... pen.” [18]	[11]
Tropological	Figurative language: metaphor, simile, hyperbole, irony, narrative analogies (following	Ideological framing: mapping complex ideas onto familiar domains (e.g. politics-as-war); Emotionalization: vivid metaphor to arouse feelings; Simplification:	“Move at a glacial pace.” – metaphorical frame to ridicule and control [8]	[14]; [10]

Device group	Description	Typical Manipulative Functions	Film example (short)	Key source
	Lakoff/Johnson [14]).	one-line hyperbole to foreclose debate.		

Illustrative Analyses (Film Examples)

Below are selected examples from the four films, organized to show how specific devices function in context. Each example is tagged with the film and an

approximate scene time (hh:mm:ss). Quoted dialogue is from subtitles or scripts (marked “[trans.]” if not from an official source). Functions are interpreted based on context analysis.

Film	Device category	Extract	Manipulative function
“Wag the Dog” [15]	Lexical	<i>“The war isn’t over until I say it’s over.”</i>	Framing: Asserting definition power over reality
“Wag the Dog” [15]	Prosodic	<i>“This. Is. Nothing.”</i>	Authority: Confident rhythm minimizing crisis
“Wag the Dog” [15]	Kinetic (props)	(Shot of actor showing film reel) <i>“We need a hero. We forgot one.”</i>	Legitimation: Visual emphasis on “hero” narrative
“Wag the Dog” [15]	Tropological	<i>“The Japanese. In the caves.”</i>	Emotionalization: War metaphor, spur action
“Thank You for Smoking” [17]	Lexical	<i>“Are you familiar with the term ‘B.S.’? ... Bullshit.”</i>	Reframing: Defining scandalous term as normal arg.
“Thank You for Smoking” [17]	Prosodic	<i>“If you argue correctly, you’re never wrong.”</i>	Closure: Aphoristic tone that ends debate
“Thank You for Smoking” [17]	Kinetic	(Father standing authoritatively) <i>“I’m his father.”</i>	Authority display: Role claim to override dissent
“Thank You for Smoking” [17]	Tropological	<i>“I see a lot of white flags.”</i>	Social pressure: Surrender metaphor to shame others
“Devil Wears Prada” [8]	Lexical	<i>“Details of your incompetence do not interest me.”</i>	Legitimation: Insult framed as business critique

Film	Device category	Extract	Manipulative function
“Devil Wears Prada” [8]	Prosodic	<i>“I don’t want dacquoise... I want tortes.”</i>	Control: Firm alternatives stressed evenly
“Devil Wears Prada” [8]	Kinetic (deixis)	<i>“Where are the belts for this dress?”</i>	Salience: Object-deictic focus, demand compliance
“Devil Wears Prada” [8]	Tropological	<i>“Move at a glacial pace.”</i>	Ridicule/Pressure: Metaphor to intensify insult
“The Wolf of Wall Street” [18]	Lexical	<i>“Sell me this... pen.”</i>	Directive framing: Contest under guise of task
“The Wolf of Wall Street” [18]	Prosodic/Timing	<i>“I don’t have a pen.” – “Exactly.”</i>	Punchline: Cadence turn, tension-then-solution
“The Wolf of Wall Street” [18]	Kinetic (object)	<i>(Jordan puts out an imaginary pen) “Write your name down.”</i>	Need-construction: Embodied demonstration of demand
“The Wolf of Wall Street” [18]	Tropological	<i>“Supply and demand, my friend.”</i>	Legitimation: Economic metaphor to justify rule

Table 2. Film examples illustrating device categories and functions (quotes are approximate). Timecodes are from standard releases and may vary; “trans.” indicates subtitle/script transcription.

From these examples, some cross-cutting observations emerged:

Interdependence of devices. Manipulative force rarely lay in one channel alone. For instance, Miranda Priestly’s command “Details of your incompetence do not interest me” [8] was effective because of both her sarcastic wording and her calm, measured delivery (table 2 entries). Similarly, Jordan Belfort’s “Sell me this pen” [18] was loaded not just by words but by the physical act (handing the pen) and timing. This supports a multimodal account: lexical framing gains impact when paired with reinforcing cues [16, 449].

Prosody as force calibrator. Lines that might be contestable became authoritative through tone. In Wag the Dog, repeating “This. Is. Nothing.” with emphatic pauses (all caps indicates a lowered, resolute voice) positioned the speaker’s reassurance as

unquestionable truth. In “Thank You for Smoking”, the calm aphorism structure allowed an absurd rule (“never wrong if argued correctly”) to be taken as reasonable guidance. Prosodic cues determined whether an assertion was heard as optional advice or as a de facto command [3, 3–5].

Kinetic emphasis and compliance pressure. Embodied actions often anchored verbal content. “The Wolf-of-Wall-Street” pen scene [18] is a classic example: the empty hand and the physical gesture of writing made the abstract rule concrete. In Prada, pointing out objects (the belt, the pen) transferred focus onto a tangible item, which functions to coerce (answer the question) without additional argument. These findings are in line with Kendon’s view of gesture as integrated action [11,11–12].

Tropes as cognitive shortcuts. Metaphors and hyperboles carried ideological weight succinctly.

“Glacial pace” [8] instantly communicated contempt and institutional dominance without elaboration. “Supply and demand, my friend” [18] invoked a familiar economic worldview to foreclose dissent. Such tropes relied on shared cultural scripts. Crucially, the analysis confirms that in manipulation, tropes serve to sidestep direct confrontation and guide inference [19, 372–376].

Audience positioning. Film discourse inherently positioned the viewer as privy to manipulative mechanics. For example, “Thank You for Smoking” often broke the fourth wall (Nick Naylor addressing the camera), turning the audience into conspirators. This aligns with Dynel’s claim that viewers are ratified listeners [6,1629–1632]. In effect, the film’s own narrative technique served as a meta-device illustrating manipulation, but this study focused on in-scene cues rather than narration.

CONCLUSION

The study provided a systematic account of how film dialogue implements manipulative influence via multimodal means. An analytical flowchart and tables were produced as tools for replicating the approach. Key conclusions include: (1) Lexical reframing is more effective when paired with supporting visual/proxemic cues; (2) Prosody is a crucial “force filter” that can render orders sound normal and hypothetical frames sound factual; (3) Gestures and object cues concretize abstract pressure; (4) Figurative language often preempts argument by mapping issues into everyday schemas.

These findings underscore that film dialogue does not manipulate solely through script but through a package of signals. For discourse analysts, this suggests adapting coding schemes to account for voice and body language, not just text. For media literacy educators, the results highlight the need to teach critical awareness of non-verbal persuasion tactics.

The study has provided a systematic account of how film dialogue constructs manipulative influence through multimodal means. By combining linguistic analysis with attention to voice, gesture, and visual context, it demonstrates that manipulation in cinema is not confined to the script itself, but emerges from the interaction of several communicative layers. The analytical flowchart and summary tables developed in this research were designed not only to describe the data, but also to serve as practical tools for replicating the approach in further studies of film or other audiovisual media.

The findings lead to several important conclusions. First, lexical reframing becomes significantly more persuasive when supported by visual or proxemic cues.

A strategic choice of words—such as redefining responsibility, urgency, or moral necessity—gains additional force when reinforced by spatial positioning, camera focus, or controlled physical distance between characters. This indicates that verbal manipulation is often strengthened by subtle visual alignment.

Second, prosody functions as a crucial “force filter.” Through changes in intonation, rhythm, stress, and volume, a direct order can be softened into what sounds like a suggestion, while a hypothetical statement may be delivered with such certainty that it is perceived as fact. In this sense, vocal delivery shapes how utterances are interpreted, often masking coercion under a tone of normalcy or inevitability.

Third, gestures and object-related cues play an important role in concretizing abstract pressure. Movements such as pointing, stepping closer, blocking space, or interacting with symbolic objects make implicit power dynamics physically visible. These non-verbal elements allow manipulation to operate without explicit threats or commands.

Finally, figurative language often works by pre-structuring interpretation. By mapping complex situations onto familiar everyday schemas, metaphors and tropes reduce the space for alternative viewpoints and subtly guide the audience toward a preferred reading. In doing so, they can preempt argument and minimize resistance.

Overall, the findings confirm that film dialogue manipulates not through isolated words, but through a coordinated package of linguistic, vocal, and visual signals. For discourse analysts, this highlights the need to adapt coding models so they account for voice and body language alongside textual analysis. For media literacy education, the findings emphasize the importance of teaching audiences to recognize not only verbal persuasion, but also the non-verbal techniques that make manipulative influence appear natural and unremarkable.

REFERENCES

1. Allan, K. & Burridge, K. (1991). *Euphemism & Dysphemism: Language Used as Shield and Weapon*. Oxford Univ. Press.
2. Bateman, J. A. & Schmidt, K.-H. (2012). *Multimodal Film Analysis: How Films Mean*. Routledge. [<https://doi.org/10.4324/9780203128220>]
3. Brazil, D. (1997). *The Communicative Value of Intonation in English*. Cambridge Univ. Press.
4. Brown, P. & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge Univ. Press.
5. Crespo-Fernández, E. (2014). *Euphemisms and*

- political discourse in the British regional press. *Brno Studies in English*, 40(1), 5–26. [https://doi.org/10.5817/BSE2014-1-1]
6. Dynel, M. (2011). "You talking to me?" The viewer as a ratified listener to film discourse. *Journal of Pragmatics*, 43(6), 1628–1644. [https://doi.org/10.1016/j.pragma.2010.11.016]
 7. Fairclough, N. (2001). *Language and Power* (2nd ed.). Longman.
 8. Frankel, D. (Dir.) (2006). «Devil Wears Prada» [Film]. 20th Century Fox.
 9. Guellaï, B., Langus, A. & Nespors, M. (2014). Prosody in the hands of the speaker. *Frontiers in Psychology*, 5, Article 700. [https://doi.org/10.3389/fpsyg.2014.00700]
 10. Kačmárová, A. (2014). Exploring film dialogue discourse. *Discourse and Interaction*, 7(1), 67–77. [https://doi.org/10.5817/DI2014-1-67]
 11. Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge Univ. Press.
 12. Kozloff, S. (2000). *Overhearing Film Dialogue*. Univ. of California Press.
 13. Kress, G. & van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design* (2nd ed.). Routledge.
 14. Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*. Univ. of Chicago Press.
 15. Levinson, B. (Dir.) (1997). "Wag the Dog" [Film]. New Line Cinema.
 16. O'Halloran, K. L. (2008). Systemic functional–multimodal discourse analysis: Constructing ideational meaning using language and imagery. *Visual Communication*, 7(4), 443–475. [https://doi.org/10.1177/1470357208096210]
 17. Reitman, J. (Dir.) (2005). "Thank You for Smoking" [Film]. Fox Searchlight Pictures.
 18. Scorsese, M. (Dir.) (2013). "The Wolf of Wall Street" [Film]. Paramount Pictures.
 19. van Dijk, T. A. (2006). Discourse and manipulation. *Discourse & Society*, 17(3), 359–383. [https://doi.org/10.1177/0957926506060250]