***Lingua: Jurnal Ilmiah**, Vol.22 No.2 October, 2026*
e-ISSN: 2715-6478

More than bad words: Sociopragmatic functions of taboo language as a communicative resource in video game dialogue

I Made Astu Mahayana[1], Luh Ketut Indrayani Pratiwi[1], Gusti Ketut Ngurah Purnama Putra[1], Emmanuel Chinedu Elele[2]

1Faculty of Literary Studies, Universitas Warmadewa, Denpasar, Indonesia
2Faculty of Education, Universitas Pendidikan Ganesha, Singaraja, Indonesia
**Corresponding author: astumahayana@gmail.com*

ABSTRACT

Taboo language performs emotional, relational, and linguistic functions in mediated discourse, but its use in scripted video game dialogue remains comparatively underexplored. This study examined the types and sociopragmatic functions of taboo language in the mission dialogue of *Grand Theft Auto: San Andreas*. Using qualitative content analysis supported by descriptive frequency counts, the study identified 1,039 taboo-language instances from 68 purposively selected narrative missions. The dialogue was recorded and transcribed with information about speakers, addressees, surrounding turns, and narrative situations. Each instance was classified using Battistella's lexical-semantic typology and analyzed for psychological, social, and linguistic functions using Crystal's framework. Intercoder agreement for lexical classification was strong ($\kappa = .81$). Obscenity was the most frequent category (529; 50.9%), followed by epithets (337; 32.4%), profanity (104; 10.0%), and vulgarity (69; 6.7%). Contextual analysis indicated that obscenity commonly intensified conflict, epithets contributed to interpersonal positioning, profanity communicated emotional evaluation, and vulgarity expressed affect or strengthened directives. Separating lexical classification from functional interpretation provides a clearer account of how taboo expressions operate in scripted game dialogue and extends research on mediated swearing to video game discourse.

Keywords: digital discourse; Grand Theft Auto: San Andreas; sociopragmatics; taboo language; video game dialogue.

Received:17 May 2026 *Revised:30 August 2026* *Published online: 20 October 2026*

INTRODUCTION

Swearing is a recurrent feature of everyday and mediated communication. Far from serving only as verbal aggression or signaling linguistic deficiency, taboo expressions can externalize emotion, intensify propositions, construct identity, negotiate relationships, and threaten another person's social standing (Bednarek, 2019; Stapleton et al., 2022). These meanings do not reside in lexical form alone. They emerge from the circumstances of use: who is speaking, who is addressed, how the participants are related, and what the interaction requires at that moment. Corpus-pragmatic research similarly indicates that the force of taboo language

varies with discourse position, interpersonal setting, and communicative purpose (Love & Stenström, 2023).

Such context sensitivity makes taboo language a sociopragmatic phenomenon. Culpeper and Tantucci's (2021) principle of (im)politeness reciprocity describes how participants adjust their relational conduct in response to one another's behavior. A potentially offensive expression may therefore convey hostility, familiarity, emotional reaction, or emphasis depending on its interactional setting. Interpretation requires attention to the surrounding turns, the expression's target, the speakers' established relationship, and the addressee's response. This approach is consistent with research identifying participants, setting, turn organization, and nonverbal behavior as important dimensions of conversational meaning (Almekky et al., 2023). It also points to a necessary analytical distinction between the lexical character of a taboo expression and the communicative work it performs.

Two complementary frameworks provide the basis for that distinction. At the lexical-semantic level, Battistella's (2005) typology differentiates epithets, profanity, vulgarity, and obscenity. Epithets are directed toward people or groups, often invoking identity, status, or personal characteristics, whereas profanity draws sacred or religious language into secular or inappropriate use. Vulgarity refers to the body and bodily processes, while obscenity encompasses strongly offensive expressions with a principally sexual basis. These categories identify the type of taboo expression used without presuming what it accomplishes in a particular interaction.

Crystal's (2018) framework addresses communicative function rather than lexical type. Psychological functions involve the externalization of emotions such as anger, fear, frustration, or surprise. Social functions concern the management of interpersonal relationships, identity, distance, solidarity, and power. Linguistic functions include intensification and other ways in which taboo expressions strengthen the communicative force of an utterance. These functions may overlap: one expression can communicate emotion while intensifying a statement or positioning another participant socially. Keeping the frameworks analytically separate therefore allows lexical classification to precede, rather than be determined by, functional interpretation.

Scripted language offers a relevant setting in which to examine this distinction, although fictional dialogue should not be treated as a direct reproduction of spontaneous conversation. Research on television discourse shows that taboo expressions contribute to emotional representation, characterization, interpersonal relationships, and metapragmatic negotiation (Bednarek, 2019; Fägersten & Bednarek, 2022). Pragmatic analysis of animated fictional dialogue has likewise examined how directive and expressive speech acts contribute to character interaction and communicative purpose (Wardana et al., 2023). Such studies support the analysis of scripted discourse while recognizing that its

linguistic features are deliberately selected, performed, and positioned within a narrative.

Video game dialogue presents a related but distinctive form of mediated discourse. Games combine scripted verbal interaction with visual representation, player action, and narrative progression. Wildfeuer and Stamenković (2022) show how meaning in games emerges through the interaction of verbal, visual, and ludic resources. Balogh and Veszelszki (2020), meanwhile, demonstrate the relevance of interpersonal language through their examination of politeness and insult in computer games. The development of the Video Game Dialogue Corpus, which contains more than 6.2 million words from 50 role-playing games, further indicates that game dialogue can be collected and analyzed systematically (Rennick & Roberts, 2024).

Despite these developments, much of the available research on video game language has focused on multimodal organization, corpus construction, linguistic representation, politeness, localization, or communication among players. Less attention has been given to how different types of taboo language perform psychological, social, and linguistic functions in scripted mission dialogue. Consequently, it remains unclear how lexical categories of taboo language relate to emotion, interpersonal positioning, conflict, and intensification within specific game interactions. The present study addresses this gap by combining lexical-semantic classification with context-sensitive functional analysis to identify both the types of taboo language used and what they accomplish within particular exchanges.

Grand Theft Auto: San Andreas (Rockstar North, 2004) provides the data for this study. Its narrative presents loyalty, rivalry, betrayal, authority, and group identity across varied interpersonal situations. The mission dialogue includes familiar exchanges among associates as well as commands, threats, and confrontations between adversaries. Because the script is fixed, the same exchanges can be recorded, transcribed, and examined consistently. The study does not treat the game as representative of all video games or as a form of natural conversation. Instead, it examines the dialogue as constructed fictional discourse in which language contributes to characterization, relationships, and narrative development.

Accordingly, the study aims to identify and describe the distribution of taboo-language types in the selected mission dialogue and to examine their psychological, social, and linguistic functions within specific interactional and narrative contexts. These objectives are addressed through the following research questions:

1. What types of taboo language occur in the selected scripted mission dialogue of Grand Theft Auto: San Andreas?
2. What psychological, social, and linguistic functions do these expressions perform within their interactional and narrative contexts?

By addressing these questions, the study adds a context-based account of taboo language to research on mediated discourse. Its main analytical contribution is the separation of lexical-semantic classification from context-sensitive functional interpretation. The study also provides a documented case that may support future comparisons across games and other forms of fictional dialogue.

METHOD

This study used qualitative content analysis supported by descriptive frequency counts. Qualitative content analysis permits systematic classification while preserving the contextual interpretation of textual data (Lindgren et al., 2020). Qualitative interpretation remained primary because the analysis considered speaker identity, interpersonal relationships, surrounding turns, and narrative circumstances. Frequencies and percentages were used only to summarize the distribution of lexical categories, not for statistical inference or hypothesis testing. The study examined pre-existing fictional dialogue and did not involve human participants.

Data were obtained from the English-language PC edition of *Grand Theft Auto: San Andreas* (Rockstar North, 2004), using the original, non-remastered Version 1.0 without modifications affecting dialogue, subtitles, missions, or character interactions. The corpus comprised 68 purposively selected narrative missions from the Los Santos, Badlands, San Fierro, Desert, and Las Venturas sections of the game. Purposive selection was used to align the corpus with the study's analytical requirements (Campbell et al., 2020). Missions were selected when they contained sustained scripted exchanges involving named characters and provided sufficient sequential and narrative context for sociopragmatic analysis. Mission cutscenes and mission-linked exchanges were included. Ambient speech, repetitive pedestrian dialogue, procedural radio announcements, speech from unnamed background characters, player-generated language, generic service activities, and dialogue repeated through failed or replayed missions were excluded. Supplementary Table S1 provides a complete inventory of the 68 missions.

Data collection took place from February 15 to March 22, 2026. The selected missions were played and recorded between February 15 and March 10. Dialogue transcription and identification of candidate taboo expressions were undertaken concurrently and completed on March 21, while final checks for eligibility and duplicate entries were completed on March 22. Candidate expressions were entered into a structured coding sheet. Each entry included the complete utterance, speaker, addressee or target where applicable, mission and scene, surrounding conversational turns, and narrative context. Screenshots and playthrough recordings were used to verify the dialogue and its context. The final dataset contained 1,039 taboo-language instances. One occurrence of a taboo lexical item or fixed expression within an utterance constituted one unit of analysis. Multiple taboo tokens in the same utterance were recorded separately.

Coding proceeded in two stages. First, each instance was assigned to one of Battistella's (2005) lexical-semantic categories: epithets, profanity, vulgarity, or obscenity. Classification was based on the expression's conventional semantic domain and contextual meaning rather than its communicative function. When an expression appeared compatible with more than one category, its meaning within the complete utterance was compared with the operational definitions before one category was assigned. These categories were mutually exclusive for frequency reporting.

Second, each instance was analyzed using Crystal's (2018) psychological, social, and linguistic functions. Functional codes were non-exclusive because an expression could perform more than one function. Interpretation considered the participants' relationship, the expression's target, the speech act, narrative circumstances, the addressee's response, and the subsequent turn. This two-stage procedure prevented communicative function from determining lexical-semantic classification.

To assess the consistency of lexical-semantic classification, a second coder independently classified a randomly selected 12% subsample of approximately 125 instances using the same operational definitions of Battistella's categories. The reliability assessment was limited to the four mutually exclusive lexical categories and did not include Crystal's overlapping functional codes. Following guidance on intercoder reliability in qualitative research (O'Connor & Joffe, 2020), agreement was calculated using Cohen's kappa. The resulting value of $\kappa = .81$ indicated strong intercoder agreement. Disagreements were resolved through discussion and reconsideration of the operational definitions

RESULTS AND DISCUSSION

The results are organized around the two research questions. Table 1 presents the distribution of the four mutually exclusive lexical-semantic categories. The subsections that follow integrate these findings with contextual interpretation and relevant literature. The frequencies describe the selected mission corpus only; they neither establish statistical differences nor represent the game's complete dialogue. Because functional codes could overlap, percentages are reported for lexical categories but not for psychological, social, and linguistic functions.

Table 1. Distribution of taboo-language types in the selected mission corpus

Type	Frequency	Percentage
Obscenity	529	50.9%
Epithets	337	32.4%
Profanity	104	10.0%
Vulgarity	69	6.7%
Total	**1,039**	**100%**

As shown in Table 1, obscenity constituted slightly more than half of the 1,039 instances, while epithets formed the second-largest category. Profanity and

vulgarity occurred less frequently. These results answer the first research question by showing the relative prevalence of the four lexical categories within the selected corpus. They do not, however, explain what the expressions accomplished in interaction. Functional interpretation required attention to speakers, targets, surrounding turns, grammatical position, and narrative circumstances.

Table 2 presents representative extracts to show how lexical-semantic classification was distinguished from sociopragmatic interpretation. Each example includes the mission, speaker, relevant dialogue or narrative circumstance, lexical basis, function, and contextual evidence.

Table 2. Representative taboo-language instances and their sociopragmatic functions

Category	Extract and context	Lexical basis	Function(s)	Contextual evidence
Epithet	**Mission: *Big Smoke***. Big Smoke: "You picked the wrong house, fool!" CJ: "Hey, hey! Smoke, it's me, Carl! Chill, chill!"	Person-directed derogatory label	Social: aggression and distancing	Big Smoke mistakes CJ for an intruder, and the expression directly targets CJ during the initial confrontation.
Profanity	**Mission: *Drive-By***. Sweet: "Those damn Ballas are rolling up on our block!"	Religious concept of damnation used in a secular expression	Psychological and linguistic: negative evaluation and intensification	The expression communicates alarm and strengthens Sweet's negative evaluation of the approaching threat.
Vulgarity	**Mission: *Running Dog***. CJ: "Oh, shit!" *(Reacting to the sudden appearance of Ballas gang members.)*	Excretory or bodily reference	Psychological: immediate emotional reaction	The expression follows the appearance of a sudden threat and evaluates the situation rather than targeting another participant.
Vulgarity	**Mission: *Life's a Beach***. OG Loc: "Man, I need that sound system, CJ. Get your ass down there and help me out."	Body-related expression	Linguistic: intensification; social: directive force	The taboo element strengthens the imperative and contributes to the forceful register of OG Loc's request.
Obscenity	**Mission: *Drive-By***. CJ: "Man, fuck those guys, let's get 'em."	Strongly offensive expression with a sexual basis	Psychological, social, and linguistic: anger, opposition, and intensification	The expression targets an opposing group during an impending confrontation and intensifies CJ's commitment to retaliatory action.

1.1 Epithets: Aggression and Interpersonal Positioning

Epithets accounted for 337 instances, representing 32.4% of the dataset. As person- or group-directed forms, they provided a direct means of evaluating an addressee or target. Their sociopragmatic force nevertheless depended on the speakers' perceived relationship and the activity in which the expression occurred.

The extract from *Big Smoke* illustrates this dependence. *Fool* is classified as an epithet because it assigns a negative description to the addressee. Its function in this exchange is aggressive because Big Smoke believes CJ to be an intruder and uses the expression during a physical confrontation. CJ's response— "It's me, Carl!

Chill, chill!"—corrects the mistaken identity and attempts to stop the confrontation. The force of *fool* therefore arises not only from its derogatory meaning but also from Big Smoke's mistaken perception of CJ and the threatening action surrounding the utterance.

Once Big Smoke recognizes CJ, the interaction changes from confrontation to familiarity. This change does not establish that *fool* itself functions as an in-group marker. Rather, it shows how updated relational knowledge can alter the direction and tone of the exchange. The evidence supports the narrower conclusion that interpreting a person-directed taboo expression requires attention to the target, the speakers' perceived relationship, and the addressee's response.

This reading is consistent with Culpeper and Tantucci's (2021) argument that evaluations of politeness and impoliteness are shaped by participant expectations and reciprocal conduct. It also corresponds with Bednarek's (2019) account of taboo language as a resource for characterization and interpersonal positioning in fictional discourse. Within the present corpus, epithets made social alignment and opposition more explicit, but their interactional significance could not be determined from lexical form alone

1.2 Profanity: Immediate Evaluation and Affect

Profanity accounted for 104 instances, or 10.0% of the dataset. These expressions drew on sacred or religious concepts used in secular contexts. In the analyzed exchanges, they commonly communicated surprise, alarm, frustration, or negative evaluation and were often directed toward a situation rather than a person.

In the *Drive-By* extract, *damn* draws on the religious concept of damnation but is used secularly to strengthen *Sweet's* evaluation of the approaching Ballas. Psychologically, it communicates alarm and hostility toward the developing threat. Linguistically, it intensifies the utterance. Removing the expression would leave the proposition— "Those Ballas are rolling up on our block"—intelligible, but its emotional and evaluative force would be reduced.

The example also shows why lexical category and communicative function were coded separately. *Damn* was classified as profanity because of its lexical-semantic basis, not because it expressed emotion. Its psychological and linguistic functions became apparent from its position in the utterance and the threat represented in the scene. This distinction prevents emotional effect from being treated as a defining feature of profanity itself.

The finding is consistent with Stapleton et al.'s (2022) account of swearing as a resource for communicating emotional intensity. In this scripted setting, the expression also contributes to the pacing of the scene by condensing evaluation and emotional response into a short form. This interpretation concerns its use within the selected mission exchange and does not imply that profanity necessarily performs the same role across contexts.

1.3 Vulgarity: Exclamation and Directive Force

Vulgarity was the least frequent category, with 69 instances representing 6.7% of the dataset. These expressions referred to the body or bodily processes, but their functions varied with grammatical form, target, and position in the interaction. The two examples in Table 2 provide a useful contrast between a situational exclamation and a person-directed instruction.

In the *Running Dog* extract, *shit* is classified as vulgarity because of its excretory meaning. In this context, it performs a psychological function: the exclamation follows the sudden appearance of a threat and externalizes CJ's immediate reaction. It evaluates the situation rather than attacking another participant. Its short, independent form also allows emotion to be communicated without explaining the speaker's state directly.

A different pattern appears in the *Life's a Beach* extract. In "Get your ass down there and help me out," *ass* occurs within an imperative directed at CJ. Removing the taboo element would not change the central instruction, but it would reduce the forceful register of the utterance. The expression therefore performs a linguistic function by intensifying the directive and a social function by shaping the position assigned to the addressee.

The contrast is analytically important because both expressions belong to the same lexical category but operate differently. In *Running Dog*, vulgarity functions as a situational reaction; in *Life's a Beach*, it modifies a directive. The difference is explained by syntax, target, and sequential context rather than by lexical category. Crystal's (2018) framework accommodates this variation because psychological, social, and linguistic functions can overlap without altering lexical classification

1.4 Obscenity: Conflict and Narrative Intensity

Obscenity was the most frequent category, comprising 529 instances, or 50.9% of the dataset. Obscene expressions appeared repeatedly in scenes involving conflict, threat, hostility, and heightened narrative tension. Their prominence should be interpreted in relation to the conflict-oriented missions selected for analysis rather than as a general pattern of video game language.

In the *Drive-By* extract, *fuck* is classified as obscenity because of its strongly offensive sexual basis. CJ uses the expression to evaluate an opposing group immediately before retaliatory action. Removing it would leave the basic intention—opposition to the group and willingness to confront it—recoverable, but the utterance would lose much of its emotional and interpersonal force.

The expression performs several functions at once. Psychologically, it externalizes anger. Socially, it marks the Ballas as an opposed group. Linguistically, it intensifies CJ's commitment to the proposed action. These functions arise from

different features of the same utterance: the speaker's emotional state, the targeted group relationship, and the strengthening of an otherwise recoverable proposition. The example therefore illustrates multifunctionality without assigning the token to more than one lexical category.

The high frequency of obscenity does not show that the game reproduces the distribution of swearing in spontaneous conversation. Scripted dialogue is selected and arranged for narrative purposes, allowing emotionally marked language to be concentrated in confrontations and other consequential scenes. In this corpus, obscenity contributes to the representation of anger and opposition while also increasing the perceived intensity of conflict. This interpretation is consistent with research showing that taboo expressions in mediated discourse contribute to emotion, characterization, and interpersonal relationships (Bednarek, 2019; Beers Fägersten & Bednarek, 2022; Stapleton et al., 2022).

Love and Stenström's (2023) corpus-pragmatic analysis similarly shows that strongly taboo forms can develop emphatic, emotional, and idiomatic functions according to their discourse environment. The present analysis identifies comparable context dependence in scripted mission dialogue. It does not establish equivalence with spontaneous speech; instead, it shows how a scripted exchange can draw on recognizable functions of taboo language to represent conflict and emotional intensity

1.5 Multifunctionality Across Lexical Categories

Lexical classification did not determine a single sociopragmatic function. Battistella's (2005) typology identified the lexical-semantic type of each expression, while Crystal's (2018) framework supported interpretation of what the expression accomplished in context. Profanity combined emotional evaluation with intensification; vulgarity functioned either as an immediate reaction or as reinforcement of a directive; and obscenity communicated anger, opposition, and intensified commitment to action.

This pattern explains why coding proceeded in two stages. Lexical categories were mutually exclusive because each token contributed once to the frequency distribution. Functional codes were non-exclusive because psychological, social, and linguistic effects could occur together. The procedure therefore distinguished the composition of the corpus from the interactional significance of particular expressions.

Frequency and function answer different questions. The distribution indicates which lexical categories were more common in the selected missions, but it cannot explain why an expression carried a particular force. Contextual interpretation can explain individual instances, but it does not establish how often each function occurred unless functional codes are counted separately. Maintaining this distinction reduces the risk of treating numerical prevalence as evidence of communicative meaning.

The findings broadly align with accounts of swearing as multifunctional and context-dependent (Bednarek, 2019; Love & Stenström, 2023; Stapleton et al., 2022). Expressions within the same lexical category may perform different functions because they occur in different speech acts, relationships, and narrative circumstances. Expressions from different categories may also perform similar functions, particularly affective expression or intensification. The relationship between lexical category and function is therefore context-sensitive rather than fixed.

Direct frequency comparisons with natural conversation would be inappropriate because the corpus was purposively constructed, the reported values were not normalized by total word count, and scripted dialogue is produced under different conditions from spontaneous speech. The contribution of these findings lies in identifying how lexical types and sociopragmatic functions relate within selected mission exchanges, not in claiming that the game reproduces natural patterns of swearing.

1.6 Target and Position in the Utterance

Comparison of the representative extracts shows that communicative function depends partly on what an expression targets and where it occurs within an utterance. This attention to target and sequential position is consistent with the sociopragmatic view that meaning develops through participants' relationships and surrounding interaction (Culpeper & Tantucci, 2021). In *Big Smoke*, *fool* identifies CJ as the target and contributes to aggression during the mistaken confrontation. In *Drive-By*, *fuck* targets an opposing group and communicates anger while reinforcing social opposition. By contrast, the exclamation *shit* in *Running Dog* responds to an unfolding situation rather than evaluating another participant. Its immediate emotional function is therefore more prominent than any relational effect.

Grammatical position also influences how taboo language contributes to an exchange. In "Get your ass down there," the taboo element is embedded within an imperative. It increases the force of the directive without changing its central instruction. Similarly, *damn* strengthens Sweet's negative evaluation while leaving the underlying proposition understandable. In both cases, the taboo item modifies the force of a larger utterance rather than functioning as a complete message by itself.

These contrasts provide a more detailed account of multifunctionality. Lexical category identifies the semantic basis of an expression, but target, grammatical form, and sequential position help explain its communicative role. Considering these features together clarifies why expressions within the same lexical category may perform different functions and why expressions from different categories may produce similar effects. This interpretation is consistent

with accounts of swearing as both context-dependent and multifunctional (Crystal, 2018; Stapleton et al., 2022).

1.7 Scripted Dialogue as Mediated Discourse

The findings must also be understood in relation to the scripted nature of the data. Dialogue in *Grand Theft Auto: San Andreas* is authored, performed, and positioned within predetermined narrative sequences. It is not equivalent to spontaneous conversation, even though it draws on recognizable patterns of interpersonal language.

The prominence of obscenity and epithets may partly reflect the narrative situations represented in the selected missions. Gang rivalry, contested authority, interpersonal threat, and shifting alliances create frequent opportunities for aggressive and identity-directed language. The distribution is therefore best understood as a feature of this purposively selected fictional corpus rather than as a general pattern of natural conversation or video game discourse.

Narrative selection may also explain why taboo expressions appear frequently in confrontational scenes. Spontaneous conversation includes many routine and low-intensity exchanges, whereas fictional dialogue tends to foreground events that develop characterization or advance the plot. Scripted expressions can therefore be concentrated at moments of conflict, threat, or transition. This concentration reflects the narrative organization of the dialogue and should not be interpreted as evidence of equivalent frequency in everyday communication.

The fixed character of scripted dialogue nevertheless supports systematic analysis. Rennick and Roberts' (2024) development of a consistently coded video game dialogue corpus illustrates the growing use of game dialogue in corpus research. Larger corpora allow comparison across games and genres, whereas the present single-game study offers closer examination of the interactional circumstances surrounding individual expressions.

Game dialogue also differs from television dialogue in how it is encountered. Both forms use scripted language to represent emotion, relationships, and social positioning (Bednarek, 2019; Beers Fägersten & Bednarek, 2022). In games, however, fixed dialogue is embedded within player-controlled movement and action. The player does not author the analyzed lines but encounters them while progressing through an interactive narrative. This combination of scripted verbal material and player-directed progression makes game dialogue a distinct form of mediated discourse without making it spontaneous speech.

1.8 Implications for Game Dialogue Research

The main theoretical implication concerns the distinction between lexical category and sociopragmatic function. Treating them as separate analytical levels avoids assuming that one type of taboo expression must perform one communicative role. Contextual information—including the speakers' relationship,

the expression's target, the speech act, narrative circumstances, and the addressee's response—is needed to explain how an expression operates within an exchange.

Methodologically, the study shows how descriptive frequency counts can complement qualitative content analysis without making numerical prevalence the basis for interpreting meaning. Frequencies describe the lexical composition of the corpus, while contextual analysis explains how representative expressions function in interaction. Combining these forms of evidence provides a clearer account than either decontextualized counting or interpretation unsupported by examples.

For game discourse research, scripted mission dialogue provides material for examining how language contributes to affect, opposition, interpersonal force, and social positioning. The analysis also shows the importance of retaining surrounding turns and narrative information when presenting game-language data. An isolated taboo word cannot reveal whether it targets a participant, evaluates a situation, strengthens a directive, or contributes to interpersonal alignment. Presenting expressions within their interactional setting therefore makes the analysis more transparent and allows readers to evaluate the proposed interpretation.

1.9 Limitations and Future Research

Certain limitations should be considered when interpreting these findings. First, the corpus was purposively constructed from 68 selected narrative missions and does not cover the game's complete dialogue. The reported frequencies therefore describe the selected corpus rather than *Grand Theft Auto: San Andreas* as a whole. In addition, the dialogue was professionally scripted, so the findings cannot be applied directly to spontaneous communication among players. The analysis also reports raw token counts rather than frequencies normalized by the total number of words in the corpus. Comparisons with corpora of different sizes would require normalized measures, such as occurrences per thousand words.

Intercoder agreement was assessed only for Battistella's mutually exclusive lexical categories. Crystal's overlapping psychological, social, and linguistic functions were interpreted from surrounding turns and narrative circumstances but were not included in the reliability assessment. Future studies could develop separate coding-consistency procedures for these functional categories. The analysis was also limited to the English-language version of one game. Because taboo meanings vary across languages and cultures, localized versions may show different lexical choices and sociopragmatic functions.

Future research could compare different game genres and franchises, examine localized versions of the same game, and contrast scripted mission dialogue with player-generated interaction. Larger comparative studies could report normalized frequencies and investigate whether particular narrative genres tend to favor certain lexical categories or sociopragmatic functions. Such research would

help clarify which patterns are specific to *Grand Theft Auto: San Andreas* and which may occur more broadly across video game discourse.

CONCLUSION

Analysis of 1,039 taboo-language instances from selected missions in *Grand Theft Auto: San Andreas* showed that obscenity was the most frequent lexical category, followed by epithets, profanity, and vulgarity. Their functions varied with the surrounding interaction: obscenity intensified conflict, epithets contributed to aggression and interpersonal positioning, profanity communicated emotional evaluation, and vulgarity expressed affect or strengthened directives. Some expressions also performed more than one psychological, social, or linguistic function.

The study's main contribution is the distinction between lexical-semantic category and sociopragmatic function. Battistella's typology identified the type of expression, while Crystal's framework guided interpretation of what it accomplished in context. This separation allowed frequency patterns to be considered without treating lexical type as evidence of a fixed communicative function.

Scripted game dialogue can therefore provide useful material for studying how taboo language contributes to emotion, relationships, and narrative conflict. Although the findings are limited to selected missions from one game, they offer a basis for comparison across games, genres, and linguistic versions.

DECLARATIONS

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions

I Made Astu Mahayana: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Writing – Original Draft, Writing – Review & Editing.
Luh Ketut Indrayani Pratiwi: Methodology, Validation, Formal Analysis, Writing – Review & Editing.
Gusti Ketut Ngurah Purnama Putra: Investigation, Data Curation, Validation, Writing – Review & Editing.
Emmanuel Chinedu Elele: Methodology, Formal Analysis, Validation, Writing – Review & Editing.

Disclosure and Conflicts of Interest

The authors declare that they have no competing interests or conflicts of interest that could have influenced the research, results, or interpretation of the findings.

Declaration of Generative AI and AI-Assisted Technologies

The authors did not use generative artificial intelligence (AI) or AI-assisted technologies in the preparation and writing of this manuscript.
All authors have read and approved the final manuscript and consent to its submission and publication.

REFERENCES

Almekky, T. A., Amir, S., Ahmed, I., Sarkingobir, Y., & Syabani, A. A. (2023). Factors affecting the flow of conversation. *Lingua: Jurnal Ilmiah, 19*(1), 92–100. https://jurnal.universitaslia.ac.id/index.php/lingua/article/view/60

Balogh, A., & Veszelszki, Á. (2020). Politeness and insult in computer games: From a pragmatic point of view. *Acta Universitatis Sapientiae, Communicatio, 7*(1), 68–91. https://doi.org/10.2478/auscom-2020-0006

Battistella, E. L. (2005). *Bad language: Are some words better than others?* Oxford University Press.

Bednarek, M. (2019). "Don't say crap. Don't use swear words."—Negotiating the use of swear/taboo words in the narrative mass media. *Discourse, Context & Media, 29*, Article 100293. https://doi.org/10.1016/j.dcm.2019.02.002

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing, 25*(8), 652–661. https://doi.org/10.1177/1744987120927206

Crystal, D. (2018). *The Cambridge encyclopedia of the English language* (3rd ed.). Cambridge University Press.

Culpeper, J., & Tantucci, V. (2021). The principle of (im)politeness reciprocity. *Journal of Pragmatics, 175*, 146–164. https://doi.org/10.1016/j.pragma.2021.01.008

Fägersten, K. B., & Bednarek, M. (2022). The evolution of swearing in television catchphrases. *Language and Literature, 31*(2), 196–226. https://doi.org/10.1177/09639470221090371

Lindgren, B.-M., Lundman, B., & Graneheim, U. H. (2020). Abstraction and interpretation during the qualitative content analysis process. *International Journal of Nursing Studies, 108*, Article 103632. https://doi.org/10.1016/j.ijnurstu.2020.103632

Love, R., & Stenström, A.-B. (2023). Corpus-pragmatic perspectives on the contemporary weakening of *fuck*: The case of teenage British English conversation. *Journal of Pragmatics, 216*, 167–181. https://doi.org/10.1016/j.pragma.2023.08.008

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods, 19*, 1–13. https://doi.org/10.1177/1609406919899220

Rennick, S., & Roberts, S. (2024). The Video Game Dialogue Corpus. *Corpora, 19*(1), 93–106. https://doi.org/10.3366/cor.2024.0299

Rockstar North. (2004). *Grand Theft Auto: San Andreas* [Video game]. Rockstar Games.

Stapleton, K., Beers Fägersten, K., Stephens, R., & Loveday, C. (2022). The power of swearing: What we know and what we don't. *Lingua, 277*, Article 103406. https://doi.org/10.1016/j.lingua.2022.103406

Wardana, M. A. W., Rohmadi, M., & Setiawan, B. (2023). Menyelisik tindak tutur direktif dan ekspresif pada serial kartun Indonesia *Trung Tung*: Kajian pragmatik. *Lingua: Jurnal Ilmiah*, *19*(1), 39–57. https://jurnal.universitaslia.ac.id/index.php/lingua/article/view/57

Wildfeuer, J., & Stamenković, D. (2022). The discourse structure of video games: A multimodal discourse semantics approach to game tutorials. *Language & Communication, 82*, 28–51. https://doi.org/10.1016/j.langcom.2021.11.005