ARTICLE IN PRESS

International Journal of Research in Marketing xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

How discourse concentration and content valence drive engagement with conversational content

Gaetano Miceli [a,*], Ludovica Serafini [b], Ernesto Cardamone [c], Maria Antonietta Raimondo [a], Francisco Villarroel Ordenes [d], Irene Scopelliti [e]

[a] University of Calabria, Department of Business Administration and Law, Campus of Arcavacata, Cubo 3C, 87036 Rende, CS, Italy
[b] University of Salento, Department of Management and Economics, Via per Monteroni, 73100 Lecce, Italy
[c] San Raffaele Rome University, Department for the Promotion of Human Science and Quality of Life, Via di Val Cannuta 247, 00166 Rome, Italy
[d] University of Bologna, Department of Management, Via Carlo Cattaneo, 17, 47921 Rimini, Italy
[e] Bayes Business School, City St. George, University of London, 106 Bunhill Row, London EC1Y 8TZ, UK

ARTICLE INFO

Keywords:
Conversational content
Engagement
Discourse concentration
Valence
Empathy
Text analysis
Podcast

ABSTRACT

Conversational content, that is, content units organized around spoken interaction between two or more speakers, has become a prominent and economically significant form of communication across contemporary media environments, with podcasts alone reaching over 100 million monthly listeners and generating over $2 billion annual advertising revenue in the U.S. Despite the rapid growth of conversational media and formats such as podcasts and social media live-streams, little is known about what drives audience engagement with this type of content. Drawing on conversation analysis, this research demonstrates that discourse concentration, defined as the extent to which speaker contributions are concentrated among a few dominant voices versus distributed more evenly across many, interacts with content valence to shape audience engagement. We find that for positively-valenced content, a more concentrated conversational structure leads to higher engagement, whereas for negatively-valenced content, a more fragmented structure elicits higher engagement. These effects are mediated by empathy toward the speakers. We test this framework across two controlled experiments using transcript-based stimuli (Study 1) and AI-generated voice stimuli (Study 2), respectively, and the econometric analysis of 401 real podcast episodes comprising more than 42,000 speaker turns and approximately 209 hours of audio (Study 3). Together, these studies provide causal evidence for the interaction, progressively extending it from text-based to audio stimuli, and demonstrate its real-world relevance. These findings show that the same conversational structure can either enhance or reduce engagement, offering actionable insights for podcast producers, media platforms, and advertisers.

ARTICLE INFO
Article history
First received on 6 October 2025 and was under review for 5½ months.
Area Editor: Shankha Basu

* Corresponding author.
E-mail addresses: g.miceli@unical.it (G. Miceli), ludovica.serafini@unisalento.it (L. Serafini), ernesto.cardamone@uniroma5.it (E. Cardamone), ma.raimondo@unical.it (M.A. Raimondo), francisco.villarroel@unibo.it (F. Villarroel Ordenes), irene.scopelliti@city.ac.uk (I. Scopelliti).

https://doi.org/10.1016/j.ijresmar.2026.07.005
Received 6 October 2025;
Available online 7 August 2026

Please cite this article as: Gaetano Miceli et al., *International Journal of Research in Marketing*, https://doi.org/10.1016/j.ijresmar.2026.07.005

Accepting Editor: Sharon Ng

1. Introduction

Conversation permeates everyday life. Individuals routinely spend hours each day engaging in or listening to conversations at home, at work, and increasingly through digitally mediated formats (Bastos, 2020a). Across media environments, *conversational content* refers to content units organized around spoken interaction between two or more speakers, such as podcast episodes, live-streamed discussions on social media, and roundtable-style TV shows (Valentini & Kruckeberg, 2012). This form of content has become economically significant for both platforms and advertisers. For instance, spending on podcast advertising has surpassed $2 billion in the U.S. (Statista, 2025), and conversational podcasts dominate the category: 61% of the 451 top-ranked U.S. podcasts rely on interviews and in-depth reporting (Stocking et al., 2023), and 57% of more than 1500 podcasters report using interviews, co-hosts, or roundtables to structure their episodes (Paterson, 2023).

Among the different forms of conversational content, podcasts represent a particularly relevant context for understanding how properties of conversation shape audience engagement. Indeed, podcasts rely on conversation as the core of their value proposition - the conversation is the product (Lindgren, 2014), which makes the structural properties of speaker participation especially salient to listeners. This unique characteristic distinguishes podcasts from other media formats that embed conversations within broader visual or narrative frameworks, where audio is one of several competing channels. Despite this relevance, marketing research on what makes conversational content engaging, and specifically on podcasts, remains surprisingly limited (Packard & Berger, 2024).

Previous marketing research has extensively studied drivers of engagement with digital content formats in which conversations are not at the core of the value proposition (e.g., captions, images, influencer videos; Karagür et al., 2022; Overgoor et al., 2022; Cascio Rizzo et al., 2026). We argue that conversational content has a defining property not shared by other digital formats: it involves a *conversational structure* resulting from turn-taking patterns and speaker participation intensity (Ames, 2013; Stocking et al., 2023). While conversational structure has been acknowledged as a potential driver of consumer behavior (Grewal et al., 2022), it has been mainly studied in the context of service failure and recovery in customer-frontline interactions (Grewal et al., 2026; Villarroel Ordenes et al., 2025). As such, understanding what drives audience engagement with conversational podcast content represents a critical gap for both theory and practice.

To address this gap, we focus on conversational podcasts and the role of *discourse concentration*, defined as the extent to which speaking contributions within conversational content are concentrated among a few dominant speakers versus distributed more evenly across multiple speakers (Asher & Lascarides, 2013; Holler et al., 2021). To illustrate this concept, consider two episodes of a popular news podcast. In the first, a journalist guides listeners through the excitement of a major scientific discovery, with minimal input from other speakers. In the second, several voices, including civilians, soldiers, journalists, and hosts, share their experiences of an ongoing war, with contributions distributed evenly across speakers. We propose that the effect of discourse concentration on audience engagement depends critically on content valence, that is, whether the information conveyed is predominantly positive or negative.

Drawing on the conversation analysis literature (Grice, 1975; Hutchby, 2019), we theorize that individuals prefer more concentrated structures when processing positive content because such structures facilitate a more intimate connection with and induce empathy toward the speakers. Conversely, individuals prefer more fragmented structures when processing negative content because a more balanced conversation elicits more social validation and support, which in turn deepens empathy toward the speakers. To test these predictions, we conduct two controlled experiments in the context of podcasts, using transcript-based stimuli and AI-generated voice stimuli (Hildebrand et al., 2020). We corroborate these findings with a field study analyzing 401 podcast episodes, comprising more than 42,000 speaker turns and approximately 209 hours of audio.

Our work makes three contributions to research on conversational content and language. First, we respond to recent calls for language research in conversational settings (Grewal et al., 2022; Packard & Berger, 2024) by demonstrating how discourse concentration influences audience engagement with podcast episodes, which represent prototypical units of conversational content. Unlike previous work on marketing content focused on single-voice content (e.g., brand or influencer captions, images and video), we examine multi-speaker conversations that require new theoretical frameworks attentive to the structural properties of language. In doing so, we integrate insights from conversation analysis (e.g., Mast, 2002) with language research in marketing contexts (Humphreys & Wang, 2018), bringing conversational structure into models of audience engagement. Relatedly, we extend research on voice numerosity effects (Chang et al., 2023) by introducing discourse concentration, a construct that captures the distribution of speakers' contributions, and by showing that its effects are contingent on content valence.

Second, we identify empathy toward the speakers (Bagozzi & Moore, 1994; Hoffman, 1984) as an important psychological mechanism explaining why the effect of discourse concentration on engagement depends on content valence. We demonstrate that for positively-valenced content, concentrated conversational structures foster a more intimate connection with speakers, enhancing empathy toward them (Lieberman et al., 2022). Instead, for negatively-valenced content, fragmented structures activate social validation and support from multiple speakers, deepening empathy toward the speakers (Taylor, 1991).

Third, from a substantive perspective, we contribute to the limited marketing research on podcasts as cultural products (Ghasemi et al., 2026). We show that conversational structure affects engagement with podcast episodes differently depending on content valence, with meaningful real-world consequences for podcast producers, media platforms, and advertisers. Our field study indicates that increasing discourse concentration by 10% raises engagement with positive content by 3.09%, but lowers engagement with negative content by 3.95%.

In the remainder of the paper, we present our conceptual framework and hypotheses, followed by three studies. We conclude by discussing the implications of our research for marketing theory and practice.

2. Conceptual framework

This section establishes the theoretical foundations for understanding engagement with conversational content. We first review relevant research on conversation analysis. Then, we develop our conceptual framework and present our research hypotheses.

2.1. Conversation analysis

Over the last decade, *language research* has become a relevant area in the consumer behavior literature, including contributions on how semantics (word meanings), pragmatics (the interaction between content and context), and syntax (grammar rules) influence consumer responses to language-based stimuli (Humphreys & Wang, 2018). Using *automated text analysis* (Hartmann et al., 2019), researchers have analyzed the effects of language features and content in the domains of *influencer success* (e.g., Cascio Rizzo et al., 2024), *online reviews* (e.g., Alantari et al., 2022; Villarroel Ordenes et al., 2017), *sharing behavior* (e.g., Barasch & Berger, 2014), *speech acts* (Villarroel Ordenes et al., 2019), and *content virality* (Berger & Milkman, 2012). In contrast, language scholars have paid less attention to the conversational structure involving turn-taking between two or more speakers (Grewal et al., 2022). In general, by separating the text into turns associated with specific speakers, it is possible to analyze conversations between parties and identify the most engaging conversational structures (Marinova et al., 2017).

A *conversation* is a series of communicative acts in which two or more agents exchange information through coordinated processes (Gnewuch et al., 2017; Grewal et al., 2026). *Conversation analysis* suggests that interpersonal communication dynamics arise not only because of the speaker's intentions, but also through the coordinated interaction that emerges as participants take turns, respond to one another, and contribute through the interaction (Grice, 1975). Thus, conversation analysis emphasizes how the distribution of turns (i.e., *contributions*) determines asymmetry and power between speakers (Hutchby, 1996, 2019), investigating the potential "*lack of balance in floor holding*" that can occur during turn-taking (Kerbrat-Orecchioni, 2004).

Over the past few decades, conversation research has produced an important body of empirical studies that have introduced some core conversational constructs and investigated their effects. Table 1 synthesizes representative studies, across different disciplines, that treat conversation not as a neutral carrier of content but as a structured interaction whose form shapes social inference and downstream outcomes.

In social psychology, research on turn-taking has established that participation in group communication is patterned rather than random, identifying systematic patterns of *floor allocation* and *turn transitions* as measurable structural components of conversation (Parker, 1988), and showing that as *group size* increases, discussions shift from interactive dialogue to leader monologue (Fay et al., 2000). A social-identity tradition further showed that *conversational participation* can translate into perceived social influence, with conversational form activating identification-based influence processes rather than operating purely through argument content (Reid & Ng, 2000). Complementing these process accounts, Mast's meta-analysis (2002) synthesizes evidence that *speaking time* robustly increases perceived dominance and power, highlighting how uneven participation signals social hierarchy. Conversation-analytic and corpus-linguistic work has further shown that *turn order* and *floor time distribution* are highly uneven and non-random across speakers in multi-party storytelling (Rühlemann & Gries, 2015), that institutional roles produce systematic participation asymmetries in dual-host radio talkback (Ames, 2013), and that conversations from large corpora cluster into distinct *discourse types* based on communication purposes (Biber et al., 2021).

Marketing and consumer research have more recently shown that the structure and content of conversations influence consumer behavior and attitudes. Bastos (2020a) showed that conversations in experiential talks tend to be perceived as more substantive and socially connecting, enhancing listener happiness, while Bastos (2020b) formalized *conversational potential* as a driver of willingness to exert effort in purchasing products. Chang et al. (2023) documented *voice numerosity* effects, showing that messages delivered by multiple voices increase persuasion and effectiveness of crowdfunding videos and video ads, unless cognitive resources are depleted. Finally, Chaudhry et al. (2025) investigated gifting in Twitch communities, finding that recipients of gifts show higher sentiment and increased *conversation continuity* with the streamer, but reduced conversation continuity with peers.

Taken together, these studies examine the effects of conversational structure (turn order, interactivity, participation inequality), conversational value (conversational potential; experiential talk), and multiple-voice delivery (voice numerosity) on dominance perceptions, social influence, motivation, persuasion, and interactive engagement. These studies collectively suggest that audiences draw meaning not only from what is said, but also from how conversations unfold. Building on this foundation, we examine how conversational structure, and in particular discourse concentration, influences audience engagement with conversational content.

2.2. Conceptual model and hypotheses

Building upon the reviewed studies on conversation analysis that have emphasized the relevance of turn distribution in shaping conversational structure (e.g., Gonzàlez-Lloret, 2010; Hutchby, 2019), we define *discourse concentration* (vs. *fragmentation*) as the extent to which speaking contributions are concentrated among a few dominant speakers versus distributed evenly across speakers in multi-speaker conversational content (Asher & Lascarides, 2013; Holler et al., 2021; McAllister et al., 1994). In a concentrated conversation, a few dominant speakers account for most of the discourse, with limited interventions from others. Conversely, in a fragmented conversation, several speakers contribute more evenly to the discourse. Discourse concentration within a conversation is distinct from the number of speakers per se, as it reflects the relative length of each speaker's contributions, rather than how many speakers are present.

We aim to investigate the effect of discourse concentration on audience engagement with conversational content (hereafter:

Table 1
Representative studies on conversational features (Core conversational constructs are bolded).

Reference	Theoretical Background	Context of Investigation	Main DV	Main IV	Mediator(s)	Moderator(s)	Research Design	Main Findings
Parker (1988)	Social psychology, Conversation analysis	Face-to-face small-group interaction	Probability of **turn transitions**	**Speaking order**	–	**Conversational context (floor, broken floor, regain, non-floor)**	Observational study with event sequence modelling	Speaking turns follow systematic, context-sensitive patterns rather than random alternation. Conversational structure emerges from patterns of turn allocation (e. g., A-B-A), with variations due to contextual changes of the floor of conversation.
Fay et al. (2000)	Group communication	Face-to-face group discussions	**Conversational interactivity** (dialogue vs. serial monologue)	**Group size**	–	–	Experiments	As group size increases, discussions shift from interactive dialogue (with stronger influence from members with whom one interacted most) to a leader monologue (with stronger influence from the dominant leader).
Reid and Ng (2000)	Social identity theory, Conversational analysis	Interpersonal and small-group conversations	Perceived group members' social influence, group member social attractiveness	**Conversational participation** (words and interruptions)	–	Prototypicality of interruptions (in-group vs. out-group utterances)	Experiment	Conversation participation is positively correlated with perceived social influence. Turns obtained through prototypical interruptions are more strongly correlated with social influence, but not with social attraction.
Mast (2002)	Social psychology, Dominance theory	Face-to-face group interaction	Expressed, inferred, and outcome dominance	**Speaking time**	–	Gender, same (vs. different) gender groups, group size	Meta-analysis	Greater speaking time positively increases expressed and inferred dominance and power; effects are robust across contexts. The relationship between inferred dominance and speaking time is stronger for men, in same-gender groups and larger groups.
Ames (2013)	Conversation analysis, Communication Media	Radio talkback shows (dual hosts and callers)	**Turn-taking, speaking time**	**Conversational roles** (hosts vs. callers)	–	–	Qualitative analysis of broadcast interactions	Conversational roles (e.g., dual hosts vs. caller) systematically shape turn-taking and speaking time, producing asymmetric participation and structured conversational dominance in mediated settings.
Rühlemann and Gries (2015)	Discourse and conversation analysis	Spoken interaction in multi-party storytelling	Probability of **linguistic patterns** to appear in stories	**Number of N-notN-N trigrams / Total number of trigrams**	–	–	Corpus-based analysis	Turn-taking in multi-party storytelling is systematically non-random, with the Narrator-Recipient-Narrator (N-notN-N) pattern that is overrepresented in all narrative examined. Speakers differ markedly in turn frequency and distribution, producing stable participation asymmetries that shape how stories unfold.

(*continued on next page*)

Table 1 (*continued*)

Reference	Theoretical Background	Context of Investigation	Main DV	Main IV	Mediator(s)	Moderator(s)	Research Design	Main Findings
Bastos (2020a)	Conversation analysis, Interpersonal Communication	Interpersonal conversations on purchases	Listener happiness	**Conversation topic**: Experiential vs. material purchase	Substantiveness, social connection	–	Experiments	Talking about purchase experiences (vs. objects) increases listener happiness. This effect is mediated by conversation substantiveness and social connection.
Bastos (2020b)	Conversation analysis	Interpersonal conversations on purchases	Consumers' willingness to exert acquisition effort	Purchase type: Experiential vs. material	**Conversational potential**	–	Experiments	Purchases with higher conversational potential (experiential vs. material) increase consumer willingness to exert acquisition effort, because they are expected to generate more engaging and socially rewarding conversations
Biber et al. (2021)	Discourse and Conversation analysis	Large-scale spoken conversational corpora	**Discourse types** (e.g., situation-dependent commentary, joking around, conflict)	Communication purposes (e.g., figuring-things-out, sharing feelings)	–	–	Corpus-based analysis	Conversational discourses can be classified into distinct discourse types based on communication purposes, showing that conversations vary meaningfully in form, not only in content.
Chang et al. (2023)	Voice research, Persuasion	Crowdfunding videos and video ads	Crowdfunding campaign success and ad effectiveness	**Voice numerosity** (single vs. multiple voices)	Favorability of cognitive responses	Cognitive resources (e.g., rate of speech, distraction)	Field studies + experiments	Multiple voices (vs. single voice) increase attention and persuasion; this effect is mediated by favorability of cognitive responses and is reduced with lower cognitive resources.
Chaudhry et al. (2025)	Live streaming, Community Gifting	Twitch live streaming	Recipient engagement, tipping, chat sentiment, **conversation continuity**	Community gifting	–	Chat with streamer vs. peers	Natural experiment + field study with text analysis	Recipients of community gifts on Twitch are initially more likely to engage socially with both streamers and peers (e.g., additional chat messages, elevated sentiment), but are not more likely to tip. Community gifts lead then to higher sentiment and increased conversation continuity with the streamer, but reduce conversation continuity with peers.
This research	Language research, Conversational analysis	Podcast episodes	Listener engagement	**Discourse concentration**	Perceived empathy	Content valence	Experiments + field study with text and audio analysis	For positively-valenced podcast episodes, higher discourse concentration leads to higher engagement; for negatively-valenced podcast episodes, lower discourse concentration leads to higher engagement. This effect is driven by empathy.

engagement). As for social media brand content (Kanuri et al., 2024), audiences can express their engagement and interest by liking and commenting on digital content. Here, we define *engagement* as the extent to which the audience expresses liking of and interest in the conversational content episode (Robiady et al., 2021).

Prior research on multiple voices has focused primarily on the number of speakers, showing that more voices can increase consumer attention and persuasion (Chang et al., 2023). However, literature on social media conversations indicates that when more people participate in an interaction, information overload may increase (Nematzadeh et al., 2019). We depart from this focus by examining the distribution of spoken contributions across speakers (i.e., discourse concentration vs. fragmentation), rather than their count. Because concentrated and fragmented structures each carry competing advantages (increasing feelings of intimacy and closeness on one hand, Collins & Miller, 1994; Lieberman et al., 2022; more dynamic and engaging interaction and different viewpoints on the other; Chang et al., 2023), the direct effect of discourse concentration on engagement is difficult to predict.

We argue that the effect of discourse concentration on engagement depends on the content of the conversation. We define *content valence* as the extent to which conversational content features positive (e.g., an entertainment podcast) versus negative content (e.g., a podcast on true crime stories – Stocking et al., 2023). Differences in the emotional tone of conversations may alter how easily audiences emotionally connect with speakers, pointing to *empathy toward the speakers* as an important mechanism explaining how discourse concentration influences engagement.

Following Hoffman (1984), empathy is a state in which people become aware of others' internal states by vicariously adopting their perspective, becoming aware of the positive or negative experiences of others (Decety & Jackson, 2004; Wispé, 1986). When attending to positively-valenced content, a more concentrated conversational structure allows dominant speakers to develop their narratives more fully, enabling greater depth of disclosure. Marx et al. (2021) demonstrated that the more a speaker reveals personal, intimate content and positive life stories, the more the listening experience is enhanced, leading to more favorable evaluations. This depth of disclosure fosters a more intimate connection between the audience and the speakers, facilitating the transfer of positive emotions and reinforcing empathy toward the speakers (Collins & Miller, 1994), which in turn increases engagement.

Consider, for instance, "*A View of the Beginning of Time*" (The Daily, July 15, 2022), an episode in which a single journalist guides listeners through the excitement of the James Webb Space Telescope's first images, with minimal input from other speakers. The concentrated conversational structure allows listeners to form an intimate connection with that speaker, facilitating the transfer of positive emotions and enhancing empathy toward the speaker. This episode exemplifies the kind of concentrated conversational structure that, we argue, leads to higher engagement through increased empathy toward the speakers for positively-valenced content.

When attending to negatively-valenced conversational content, individuals seek social validation and support to make sense of difficult information (Baumeister et al., 2001; Taylor, 1991). A fragmented conversational structure, featuring multiple, balanced voices, provides precisely this: each speaker's contribution signals that others validate the emotional weight of the situation, creating a sense of collective acknowledgement. John et al. (2019) demonstrated that people dislike receiving negative news from a single messenger due to perceived malevolent motives. When confronted with bad news, individuals may seek additional viewpoints to make sense of the situation (Baumeister et al., 2001). Consistent with this view, research on stress responses to negative information suggests that group discussions and the exchange of different viewpoints ultimately alleviate distress (Rimé, 2009). When multiple speakers share the conversational floor in response to negative content, each contribution serves as a signal that others recognize and validate the emotional weight of the situation. Audiences perceive this distributed acknowledgment as evidence of collective emotional charge, which in turn activates empathic concern toward the speakers. Research on responsive conversational disclosure shows that when one party reveals a negative outcome, others try to provide emotional comfort and distribute the burden of negative affect; this behavior is driven by concern for the counterpart's emotional state (Prinsloo et al., 2026). Empathy toward the speakers thus emerges as the proximal psychological mechanism through which fragmented conversational structures translate social validation cues into higher audience engagement. Conversely, negatively-valenced conversations dominated by one or a few speakers may fail to foster a sense of connection and social support, thereby reducing perceived empathy toward the speakers and engagement (Luminet IV et al., 2000).

Consider the podcast episode "*In Ukraine, the Men Who Must Stay and Fight*" (The Daily, March 1, 2022). Here, fifteen speakers, including Ukrainian civilians, soldiers, journalists, researchers and hosts, share their personal experiences of the war, with contributions distributed evenly across voices. The fragmented structure signals that the weight of the negative content is collectively acknowledged across multiple voices, activating social validation and deepening listeners' empathy toward the speakers as a group. This episode exemplifies the kind of fragmented conversational structure that, we argue, leads to higher engagement through increased empathy toward the speakers for negatively-valenced content.

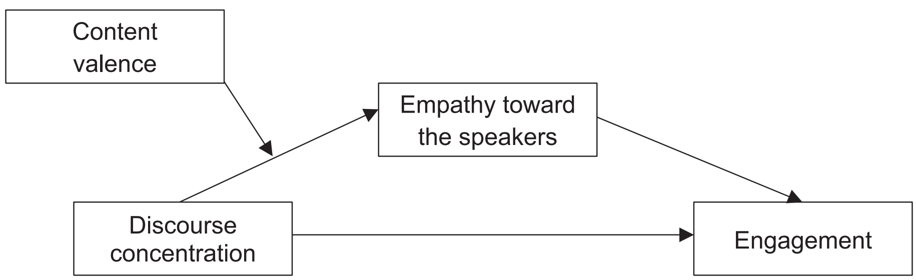


Fig. 1. A conceptual model of the effect of discourse concentration on engagement.

In sum, we propose that the effect of discourse concentration on engagement is moderated by content valence, and that this interaction effect is mediated by empathy toward the speakers. Formally:

H1. Content valence moderates the effect of discourse concentration on engagement: For more positively-valenced content, higher discourse concentration (i.e., lower fragmentation) generates higher engagement, whereas for more negatively-valenced content, lower discourse concentration (i.e., higher fragmentation) generates higher engagement.

H2. The effect of discourse concentration on engagement, moderated by valence, is mediated by empathy toward the speakers.

Fig. 1 depicts the proposed conceptual model. We test this conceptual model by means of three studies, using different research designs.

2.3. Empirical context: podcasts

We selected podcasts as an ideal multiple-speaker setting to examine engagement drivers in conversational media. Podcasts are "*episodical, downloadable or streamable, primarily spoken audio content, distributed via the Internet, playable anywhere, at any time, produced by anyone who so wishes*" (Rime et al., 2022, p. 1270). Among the different types of podcasts (Berry, 2020), *conversational podcasts* dominate the category (Stocking et al., 2023) and are the most common format (Paterson, 2023). While conversational content spans a wide range of formats, podcasts offer a setting in which conversational dynamics are *central* rather than ancillary to the content. Unlike other communication formats that embed conversations within broader narratives or visual frameworks, for podcasts the conversation *is* the product (Lindgren, 2014). Moreover, podcasts create a uniquely intimate connection between speakers and listeners (Llinares, 2018) through the immersive listening experience (Lieberman et al., 2022; Rajar, 2020). This creates what researchers call an "intimate bridging medium" (Swiatek, 2018), where the listening experience feels personal and direct rather than broadcast to a mass audience.

Beyond their theoretical distinctiveness, podcasts are also understudied in marketing research (see Ghasemi et al., 2026 for a recent exception). While prior studies have examined how structural and linguistic features of content shape audience engagement across several media formats (e.g., Behrens et al., 2025; Berger et al., 2021), podcasts have received comparatively little attention. This gap is particularly relevant given that podcasts have become a mainstream media format, reaching a substantial share of the adult population and involving repeated, habitual listening over time for purposes ranging from entertainment and education to news consumption (Pew Research Center, 2025). For these reasons, the three studies reported in the following sections focus on audience engagement with podcast episodes, using different research designs and stimuli.

3. Empirical studies

3.1. Overview of the studies

We test our conceptual framework through three complementary studies designed to progressively disentangle the structural effects of discourse concentration from modality-specific and contextual explanations, while establishing both internal and external validity. All studies focused on episodes of *The Daily*, a widely consumed news podcast produced by *The New York Times*. In addition to its popularity (approximately 3 million daily listeners; Talbot, 2024), *The Daily* represents a paradigmatic example of conversational news content in which discourse concentration varies meaningfully across episodes, as conversational dynamics naturally fluctuate with the presence of multiple speakers, including journalists, subject-matter experts, and authoritative sources.

Study 1 tests H1 and H2 through a controlled experiment that manipulates discourse concentration (vs. fragmentation) and content valence (positive vs. negative) using written transcripts of an adapted podcast episode. Our theory posits that the interaction between discourse concentration and content valence shapes engagement independently of the modality through which speaker contributions are delivered. By focusing on textual stimuli, Study 1 isolates this interaction from paralinguistic cues such as voice tone or register, providing a clean causal test of whether the interaction between discourse concentration and content valence drives the predicted effects.

Study 2 replicates the design of Study 1 but using audio stimuli. While Study 1 establishes that the interaction between discourse concentration and content valence shapes engagement at the level of text, our framework ultimately concerns *spoken conversational content*. Podcast listeners encounter this interaction not as written text but as heard speech, where the felt experience of one dominant voice versus many voices may be more immediate. Study 2 therefore tests whether the predicted interaction holds when discourse is in auditory modality, while controlling for idiosyncratic speaker characteristics through AI-generated voices. In doing so, Study 2 represents one of the first applications of AI-generated voice stimuli to the study of conversational structure and audience engagement, contributing to the growing methodological toolkit available for research on spoken conversational content (Hildebrand et al., 2020).

Finally, Study 3 tests H1 in a naturalistic setting, by analyzing transcripts from 401 real episodes of *The Daily*. While Studies 1 and 2 establish causal evidence under controlled conditions, Study 3 tests whether the interaction between discourse concentration and content valence generalizes to real podcast episodes, where structural and content features co-vary naturally. In doing so, Study 3 demonstrates that the proposed framework explains meaningful variance in real-world audience engagement, contributing substantive evidence beyond what controlled settings alone can provide.

Taken together, the three studies offer a coherent empirical package in which complementary research designs and stimuli allow us to test our conceptualization across modalities, to ensure data richness, and to demonstrate its real-world relevance (Blanchard et al., 2022; Wang et al., 2024).

3.2. Study 1: analyzing podcast conversations in transcript-format

Study 1 is an experimental study manipulating the transcript of an episode of *The Daily*.

Stimuli. We created the stimuli for a 2 (discourse concentration: concentrated vs. fragmented) × 2 (valence: positive vs. negative) between-subjects experiment by using the transcript of *The Daily* episode "*A Broadway Show Comes Back to Life*" as a baseline. The original episode featured host Michael Barbaro interviewing the cast and crew of the Broadway musical "*Six*" about their experiences during the COVID-19 pandemic. We created four versions of the transcript and retitled them more neutrally "*Broadway Shows During the Pandemic: Six - The Musical*".

We manipulated discourse concentration by redistributing speaking turns among six speakers (three female and three male). The concentrated version featured one theatre columnist dominating the conversation (71% of content), while others had minimal input. The fragmented version had all six speakers contributing roughly equally to the conversation. To make the manipulation more salient, we assigned a distinct color to each speaker and added thumbnails with their photos next to their spoken contributions. To preliminarily check the manipulation, we computed the Hirschman-Herfindahl Index (HHI) and a conversation dominance index (e.g., Henderson & Cote, 1998) based on the relative length of the speakers' turns (in terms of the number of characters per turn) as measures of discourse concentration. HHI is computed as$\sum_{s=1}^{S} p_s^2$, where p_s represents the relative contribution length for speaker s, and ranges between 1/S (high fragmentation) and 1 (high concentration). The conversation dominance index is the largest relative contribution length (p_s), with higher values indicating higher discourse concentration. The four versions of the episode transcript showed scores consistent with the intended manipulation ($\text{HHI}_{\text{Concentration\_Positive}} = 0.55$, $\text{Dominance}_{\text{Concentration\_Positive}} = 0.71$; $\text{HHI}_{\text{Concentration\_Negative}} = 0.55$, $\text{Dominance}_{\text{Concentration\_Negative}} = 0.71$; $\text{HHI}_{\text{Fragmentation\_Positive}} = 0.17$, $\text{Dominance}_{\text{Fragmentation\_Positive}} = 0.22$; $\text{HHI}_{\text{Fragmentation\_Negative}} = 0.17$, $\text{Dominance}_{\text{Fragmentation\_Negative}} = 0.22$).

To manipulate valence, we created a positive version of the transcript, emphasizing the reopening of Broadway theatres after the pandemic and conveying positive feelings of warmth and hope; and a negative version emphasizing the shutdown of theatres, conveying feelings of gloom and despair. We preliminarily checked this manipulation by computing LIWC 2022 Tone scores (Boyd et al., 2022), which range from zero (more negatively-valenced content) to 100 (more positively-valenced content). The positive versions of the transcript show a Tone score of 97.88, while the negative versions had a Tone score of 3.31. The complete set of stimuli is available in Web Appendix 1.

Participants and Procedure. Participants accessed a Qualtrics study and were randomly assigned to one of the four experimental conditions. We recruited 431 participants based in the UK through Prolific Academic, and paid each £1 for their participation. We targeted this sample size to gather at least 75 participants per condition, anticipating some attrition due to two screening criteria employed to ensure data quality. First, we restricted participation to desktop/laptop users to ensure optimal transcript readability. Second, we required participants to spend at least three minutes reading the transcript, based on pretesting that showed this timing was necessary for adequate comprehension of the conversational structure and content. Participants who proceeded before 180 seconds elapsed were automatically redirected to the end of the survey where they answered socio-demographic questions. This screening process resulted in a final sample of 315 participants (185 females, 129 males, one other; $M_{age} = 40.39$, $SD_{age} = 11.03$). Excluded participants (53 females, 63 males; $M_{age} = 35.11$, $SD_{age} = 10.19$) were on average younger, with a slight prevalence of males. To verify that exclusions were not systematic, we ran a logistic regression with exclusion status as the dependent variable and discourse concentration, content valence, and their interaction as predictors: no effect was significant (all $ps > 0.12$). An ANOVA on stimulus reading time yielded no significant effects either (all $ps > 0.14$). Based on the observed interaction effect size, a post hoc power analysis indicated that a sample of 220 participants would be sufficient to achieve 80% power. Therefore, the final sample provided adequate statistical power (0.92).

Measures. Episode engagement was measured using two items: *How much have you enjoyed reading the transcript of the episode?* (1 = *not at all*, 7 = *very much*); and *How much were you engaged while reading the podcast episode transcript?* (1 = *not at all*, 7 = *very much*). The correlation between the two items was very high ($r = 0.66$, $p < 0.001$), so we computed a mean score of engagement. We collected data on empathy toward the speakers with five 7-point Likert items adapted from Bagozzi and Moore (1994) and Schumann et al. (2014): *I felt empathy toward the speakers in the conversation*; *The speakers' feelings really matched my feelings*; *I feel the way the speakers feel about this topic*; *I really feel deep sympathy for the speakers*; and *As I read the episode transcript, I could actually feel what the speakers were describing*. Cronbach alpha was 0.92, so we computed a mean score of empathy toward the speakers.

Participants then completed two manipulation checks for discourse concentration: "*The episode reported in the transcript features six speakers. To what extent do you think that the spoken contributions of the speakers were:* 1 = *well distributed among the speakers*, 7 = *strongly concentrated upon few speakers*"; and "*The conversation between the speakers was*: 1 = *balanced*, 9 = *unbalanced*". Since the two items were highly correlated ($r = 0.69$), we computed a mean score of the discourse concentration manipulation check. Lastly, participants completed three manipulation checks for valence: "*The transcript that you have just read was mostly*: 1 = *negatively valenced*, 9 = *positively valenced*"; "*The content of the episode was* 1 = *very negative*, 9 = *very positive*"; "*Overall, the tone of the episode was* 1 = *very negative*, 9 = *very positive*". Cronbach alpha was 0.98, so we averaged the three items. Finally, participants answered a set of socio-demographic questions.

Results. The manipulations of discourse concentration and valence were successful. A 2 (discourse concentration: concentrated vs. fragmented) × 2 (valence: positive vs. negative) ANOVA on the discourse concentration manipulation check yielded only the manipulated discourse concentration effect (concentrated: $M = 5.57$, $SD = 1.54$; fragmented: $M = 3.40$, $SD = 1.40$; $F(1,311) = 171.81$,

$p < .001$, *partial* $\eta^2 = 0.356$). Neither the manipulated valence ($F(1,311) = 0.18$, $p = .668$, *partial* $\eta^2 = 0.001$) nor the interaction ($F(1,311) = 1.71$, $p = .192$, *partial* $\eta^2 = 0.005$) effects were significant. A similar analysis on the valence manipulation check yielded only the manipulated valence effect (positive: $M = 7.99$, $SD = 0.91$; negative: $M = 3.19$, $SD = 1.28$; $F(1,311) = 1475.38$, $p < .001$, *partial* $\eta^2 = 0.826$). Neither the manipulated discourse concentration ($F(1,311) = 1.36$, $p = .245$, *partial* $\eta^2 = 0.004$) nor the interaction ($F(1,311) = 0.51$, $p = .475$, *partial* $\eta^2 = 0.002$) effects were significant.

A 2 (discourse concentration: concentrated vs. fragmented) × 2 (valence: positive vs. negative) ANOVA on engagement revealed a significant discourse concentration × valence interaction ($F(1,311) = 11.51$, $p = .0008$, *partial* $\eta^2 = 0.036$). In line with H1, for positively-valenced episodes, participants showed higher engagement with a more concentrated discourse structure (concentrated-positive: $M = 3.94$, $SD = 1.60$; fragmented-positive: $M = 3.35$, $SD = 1.28$; $F(1,311) = 6.98$, $p = .009$, *partial* $\eta^2 = 0.022$). For negatively-valenced episodes, participants were more engaged with a podcast featuring a more fragmented discourse structure (concentrated-negative: $M = 3.64$, $SD = 1.42$; fragmented-negative: $M = 4.12$, $SD = 1.27$; $F(1,311) = 4.66$, $p = .033$, *partial* $\eta^2 = 0.015$). Neither the discourse concentration ($F(1,311) = 0.11$, $p = .745$, *partial* $\eta^2 = 0.000$) nor the valence main effects ($F(1,311) = 2.22$, $p = .138$, *partial* $\eta^2 = 0.007$) were significant. Fig. 2 reports these results.

A similar 2 × 2 ANOVA on perceived empathy toward the speakers revealed a significant discourse concentration × valence interaction ($F(1,311) = 14.12$, $p = .0002$, *partial* $\eta^2 = 0.043$). The discourse concentration main effect ($F(1,311) = 0.08$, $p = .777$, *partial* $\eta^2 = 0.000$) was not significant, while the valence main effect ($F(1,311) = 24.71$, $p < .001$, *partial* $\eta^2 = 0.074$) was significant, with negatively-valenced content yielding higher empathy toward the speakers (positive: $M = 3.02$, $SD = 1.41$; negative: $M = 3.80$, $SD = 1.39$). For positively-valenced episodes, participants showed higher empathy toward the speakers with a more concentrated discourse structure (concentrated-positive: $M = 3.34$, $SD = 1.54$; fragmented-positive: $M = 2.71$, $SD = 1.20$; $F(1,311) = 8.25$, $p = .004$, *partial* $\eta^2 = 0.026$). For negatively-valenced episodes, participants showed higher empathy toward the speakers with a more fragmented discourse structure (concentrated-negative: $M = 3.52$, $SD = 1.38$; fragmented-negative: $M = 4.06$, $SD = 1.36$; $F(1,311) = 5.98$, $p = .015$, *partial* $\eta^2 = 0.019$).

To test H2, we estimated a moderated mediation model (Hayes, 2022; PROCESS Model 7) in which valence (0 = negative, 1 = positive) moderated the path from discourse concentration (0 = fragmented, 1 = concentrated) to empathy toward the speakers, and empathy toward the speakers influenced engagement. In the empathy toward the speakers model, the discourse concentration × valence interaction was positive and significant ($b = 1.17$, $p = .0002$). Simple slope analysis showed that, in the positive valence condition, the effect of discourse concentration on empathy toward the speakers was positive ($b_{\text{positive}} = 0.63$, $p = .004$), whereas in the negative valence condition the same effect was negative ($b_{\text{negative}} = -0.54$, $p = .015$). In the engagement model, the effect of empathy toward the speakers was positive and significant ($b = 0.69$, $p < .001$), while the effect of discourse concentration was not significant ($b = 0.02$, $p = .84$). We then analyzed the conditional indirect effects. In the positive valence condition, the discourse concentration ➔ empathy toward the speakers ➔ engagement indirect effect was positive and significantly different from zero (IE = 0.43, 95% Bootstrap CI [0.14, 0.75]). In the negative valence condition, the same indirect effect was negative and significantly different from zero (IE = −0.37, 95% Bootstrap CI [−0.68, −0.08]). The confidence interval of the index of moderated mediation did not include zero ($\omega = 0.80$, 95% Bootstrap CI [0.37, 1.27]). Therefore, the two conditional indirect effects were significantly different from each other. These results support H2.

Discussion. The results of Study 1 support H1 and show that the valence of the content moderates the effect of discourse concentration on engagement. Specifically, when a podcast episode conveys positive content, individuals are more engaged with a more concentrated structure of the conversation among the speakers. This effect reverses when the episode conveys negative content, as individuals are more engaged with a more fragmented structure of the conversation among speakers. Consistent with H2, this effect is mediated by empathy toward the speakers. When an episode features positive content, discourse concentration enhances empathy toward the speakers, which, in turn, positively affects engagement. For negative content, the reverse occurs: discourse fragmentation enhances empathy toward the speakers by activating social validation and support from multiple speakers. Empathetic responses, in turn, increase engagement with the fragmented conversational structure.

3.3. *Study 2: analyzing podcast conversations in an AI-generated audio format*

Study 2 tested our conceptual model using audio stimuli to reproduce voice-based content. This study was pre-registered (https://aspredicted.org/ty3j-y6tp.pdf).

Stimuli. The experiment was a 2 (discourse concentration: concentrated vs. fragmented) × 2 (valence: positive vs. negative) between-subjects design. Starting with the podcast transcript used in Study 1, we changed the speaker names to fictional ones and the show title to "Morning Bits" to avoid recognition effects. We then converted the four transcript versions into 7-minute audio episodes using *Natural Reader*, a text-to-speech software that converts written text into spoken audio using AI-generated voices. The software offers a wide range of realistic, natural-sounding voices and allows adjustments to features such as speed, pitch, and volume, which is desirable for creating experimental voice stimuli (Efthymiou et al., 2024). We selected six speakers' voices (three females, three males) from those categorized as "suitable for audiobooks and podcasts", labeled as "conversational" and "natural". Natural Reader automatically adjusted vocal tone based on the mood inferred from the text and applied a consistent pause duration between speaker turns across all conditions. Overlapping speech was intentionally avoided to maintain intelligibility and experimental control. Inter-turn pause durations were kept comparable across the four audio tracks, ruling out the possibility that the observed interaction effect is driven by systematic differences in pause structure or turn-taking timing across conditions. We post-produced the audio tracks using GarageBand (which was also used to create a neutral podcast theme song) and adjusted volume levels and layering using Audacity. The four audio-tracks are available upon request from the first author. The results of an audio analysis (Hildebrand et al., 2020) of the four

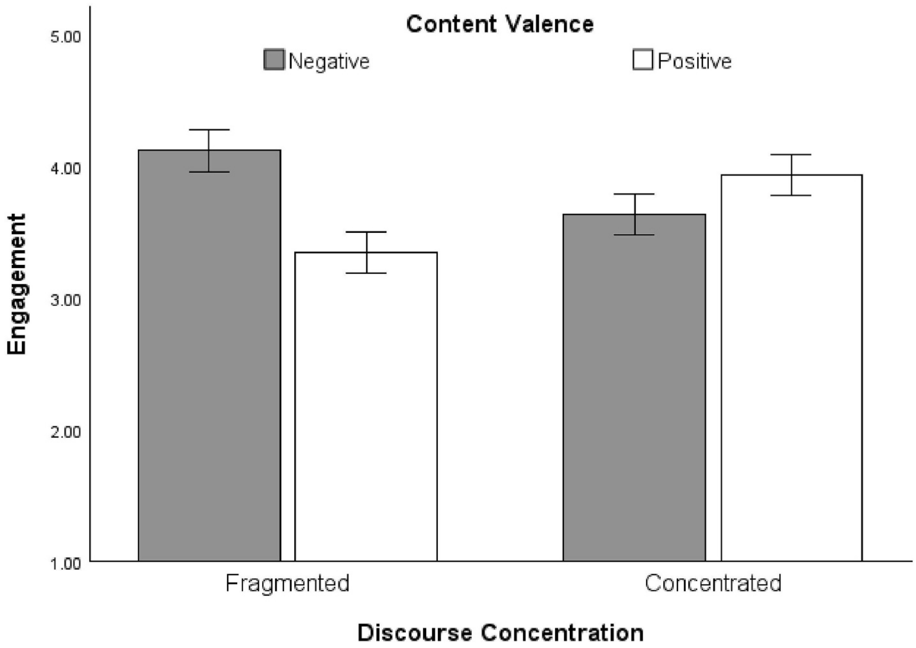


Fig. 2. Means (± 1SE) of episode engagement as a function of discourse concentration and valence (Study 1).

tracks are available in the Web Appendix 2.

To the best of our knowledge, this is one of the first studies in consumer research to employ audio stimuli generated with AI voices. To facilitate transparency and replicability, we have documented the procedure as a pipeline (see Web Appendix 3) that can be adapted for future research involving audio-based marketing stimuli (e.g., audio advertisements, branded podcasts).

Participants and Procedure. Participants accessed a Qualtrics study and were randomly assigned to one of the four experimental conditions. We recruited 430 UK-based participants through Prolific Academic, each of whom received £1 for their participation. We targeted this sample size to obtain at least 75 participants per condition, anticipating some attrition due to pre-registered screening mechanisms.

As an initial screening step, participants completed a "sound-check". They listened to a short audio clip that recited "*Mum went to the market and bought a bunch of red flowers*". Those who did not select red (from seven color options) were immediately debriefed, as they were presumed unable to properly hear the audio tracks. Ten participants failed the sound-check, leaving a sample of 420 participants (228 females, 191 males, one other; $M_{age} = 40.55$, $SD_{age} = 10.08$). Participants then listened to one of the four podcast episode versions. The button to proceed became available only at the end of the episode.

Measures. We used the same scales as in Study 1 for episode engagement ($r = 0.66$, $p < 0.001$), empathy toward the speakers ($\alpha = 0.93$), discourse concentration ($r = 0.44$, $p < .001$) and valence manipulation checks ($\alpha = 0.96$). Participants then completed two pre-registered attention checks to verify that they had actually listened to the podcast episode. Specifically, they were asked: (1) to select how many female and male speakers were involved in the episode (correct answer: 3 females, 3 males; other options: 6 males, 0 females; 6 females, 0 males; 5 males, 1 female; 5 females, 1 male); and (2) to indicate the name of the show mentioned in the episode (correct answer: *Six*; other options: *A Chorus Line*; *Hamilton*; *Mary Poppins*). Finally, participants answered socio-demographic questions.

One hundred forty-nine participants failed at least one attention check. Although this high failure rate was somewhat unexpected, we proceeded as pre-registered and analyzed data from the remaining 271 participants (150 females, 120 males, one other; $M_{age} = 41.16$, $SD_{age} = 10.23$). We ran a logistic regression with exclusion status as the dependent variable and discourse concentration, content valence, and their interaction as predictors. While the effects of valence ($b = -0.52$, $p = .13$) and the interaction ($b = 0.39$, $p = .37$) were not significant, the effect of discourse concentration was significant ($b = 0.90$, $p = .002$), suggesting that more concentrated versions led to higher probability of exclusion. While participants assigned to the concentrated conditions were more likely to be excluded overall, this tendency did not differ by valence condition. Robustness analyses on the full sample (see the *Robustness checks* subsection) confirm that the core interaction effect holds regardless of exclusions. Based on the observed interaction effect size, a post hoc power analysis indicated that a sample of 184 participants would be sufficient to achieve 80% power. Therefore, the final sample, despite the exclusions, provided adequate statistical power (0.93).

Results. The manipulations were successful. A 2 (discourse concentration: concentrated vs. fragmented) × 2 (valence: positive vs. negative) ANOVA on the discourse concentration manipulation check yielded only the manipulated discourse concentration effect (concentrated: $M = 4.63$, $SD = 1.58$; fragmented: $M = 3.43$, $SD = 1.39$; $F(1,267) = 44.814$, $p < .001$, *partial* $\eta^2 = 0.144$). Neither the manipulated valence ($F(1,267) = 0.11$, $p = .740$, *partial* $\eta^2 = 0.000$) nor the interaction ($F(1,267) = 0.85$, $p = .356$, *partial* $\eta^2 = 0.003$)

effects were significant. A similar analysis on the valence manipulation check revealed only the manipulated valence effect (positive: $M = 7.80$, $SD = 1.08$; negative: $M = 4.08$, $SD = 1.86$; $F(1,267) = 410.11$, $p < .001$, *partial* $\eta^2 = 0.606$). Neither the manipulated discourse concentration ($F(1,267) = 0.27$, $p = .605$, *partial* $\eta^2 = 0.001$) nor the interaction ($F(1,267) = 0.82$, $p = .365$, *partial* $\eta^2 = 0.003$) effects were significant.

A 2 (discourse concentration: concentrated vs. fragmented) × 2 (valence: positive vs. negative) ANOVA on engagement showed a significant discourse concentration × valence interaction ($F(1,267) = 11.62$, $p = .0008$, *partial* $\eta^2 = 0.042$). The main effects of discourse concentration (concentrated: $M = 3.86$, $SD = 1.49$; fragmented: $M = 3.52$, $SD = 1.44$; $F(1,267) = 3.05$, $p = .082$, *partial* $\eta^2 = 0.011$) and valence (positive: $M = 3.47$, $SD = 1.45$; negative: $M = 3.87$, $SD = 1.47$; $F(1,267) = 3.21$, $p = .074$, *partial* $\eta^2 = 0.012$) were marginally significant. Consistent with H1, for positively-valenced episodes, participants were more engaged with a more concentrated discourse structure (concentrated-positive: $M = 3.99$, $SD = 1.50$; fragmented-positive: $M = 3.09$, $SD = 1.29$; $F(1,267) = 14.01$, $p = .0002$, *partial* $\eta^2 = 0.050$). For negatively-valenced episodes, participants tended to be more engaged with a more fragmented discourse structure (concentrated-negative: $M = 3.71$, $SD = 1.48$; fragmented-negative: $M = 4.00$, $SD = 1.46$; $F(1,267) = 1.32$, $p = .253$, *partial* $\eta^2 = 0.005$), but this difference did not reach statistical significance. These results partially support H1. Fig. 3 synthesizes these results.

A similar 2 × 2 ANOVA on perceived empathy toward the speakers revealed a marginally significant discourse concentration × valence interaction ($F(1,267) = 3.56$, $p = .060$, *partial* $\eta^2 = 0.013$). The discourse concentration main effect ($F(1, 267) = 1.61$, $p = .206$, *partial* $\eta^2 = 0.006$) was not significant, while the valence main effect ($F(1, 267) = 12.79$, $p = .0004$, *partial* $\eta^2 = 0.046$) was significant, with negatively-valenced content yielding higher empathy toward the speakers (positive: $M = 2.98$, $SD = 1.38$; negative: $M = 3.67$, $SD = 1.50$). For positively-valenced episodes, participants showed higher empathy toward the speakers with a more concentrated discourse structure (concentrated-positive: $M = 3.31$, $SD = 1.46$; fragmented-positive: $M = 2.75$, $SD = 1.28$; $F(1,267) = 5.26$, $p = .023$, *partial* $\eta^2 = 0.019$). For negatively-valenced episodes, participants showed higher empathy toward the speakers with a podcast featuring a more fragmented discourse structure (concentrated-negative: $M = 3.60$, $SD = 1.51$; fragmented-negative: $M = 3.71$, $SD = 1.50$; $F(1,267) = 0.18$, $p = .670$, *partial* $\eta^2 = 0.001$) but this difference was not significant.

To test H2, we estimated the same moderated mediation model as in Study 1 (PROCESS Model 7). In the empathy toward the speakers model, the discourse concentration × valence interaction was positive and marginally significant ($b = 0.66$, $p = 0.060$). Simple slope analysis showed that, in the positive valence condition, discourse concentration had a positive effect on empathy toward the speakers ($b_{\text{positive}} = 0.55$, $p = .023$), whereas in the negative valence condition, the same effect was negative ($b_{\text{negative}} = -0.11$, $p = .670$) but did not reach significance. In the engagement model, the effect of empathy toward the speakers was positive and significant ($b = 0.74$, $p < .001$), while the effect of discourse concentration was not significant ($b = 0.16$, $p = .19$). We then analyzed the conditional indirect effects. In the positive valence condition, the discourse concentration ➔ empathy toward the speakers ➔ engagement indirect effect was positive and significantly different from zero (IE = 0.41, 95% Bootstrap CI [0.12, 0.71]). In the negative valence condition, the same indirect effect was negative (IE = −0.08, 95% Bootstrap CI [−0.41, 0.24]), but not significantly different from zero. The confidence interval of the index of moderated mediation does not include zero only at the 90% level (ω = 0.49, 90%

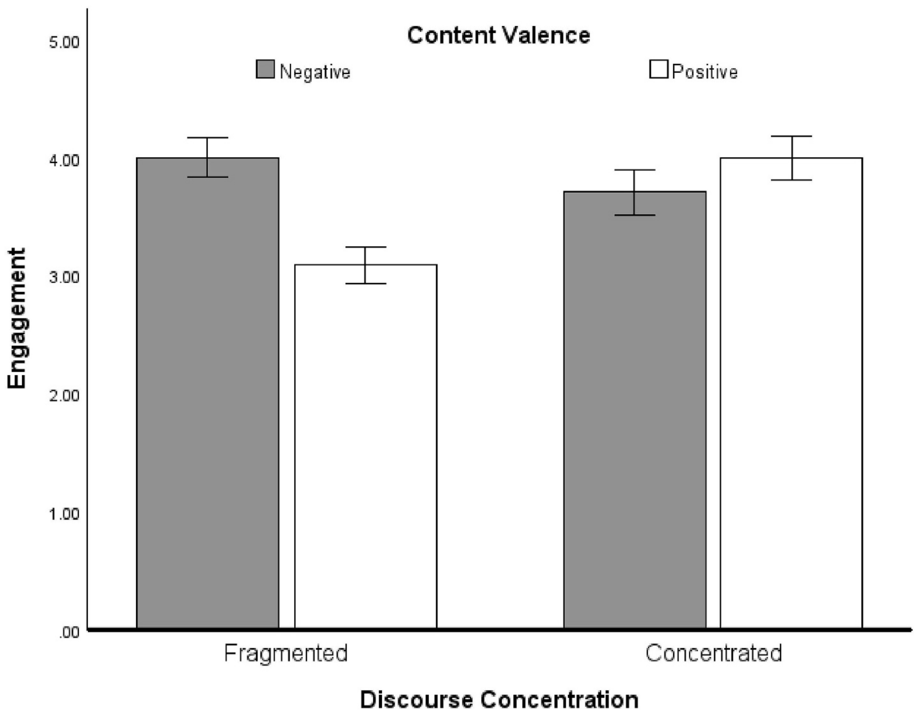


Fig. 3. Means (± 1SE) of episode engagement as a function of discourse concentration and valence (Study 2).

Bootstrap CI [0.05, 0.95]). These results offer partial support to H2.

Robustness checks. Given the unexpectedly high number of participants who failed the attention checks, we conducted some additional analyses (see Web Appendix 4) to assess the stability of the findings. In particular, we re-ran the analyses on (1) the full sample; (2) the full sample with a dummy control variable (1 = passed attention checks; 0 = failed either of them); and (3) a sample excluding only the ten participants who failed both attention checks. The results remained stable, with two differences: the interaction effect in the empathy toward the speakers model became statistically significant ($p < .05$), and the index of moderated mediation reached significance at the 95% confidence level.

Discussion. Study 2 used audio stimuli to gather additional evidence for our conceptualization. The results suggest that the effect of discourse concentration on episode engagement depends on content valence. Specifically, for positively-valenced episodes, individuals engaged more with a concentrated conversational structure. For negatively-valenced episodes, the results suggested higher engagement with more fragmented conversational structure, but this effect did not reach conventional levels of statistical significance. A similar pattern emerged for the proposed mediator, empathy toward the speakers. Overall, our hypotheses were supported using audio stimuli, but the less stable findings in the negative valence condition call for further investigation. Therefore, Study 3 examines engagement with real podcast episodes to test the predicted interaction in a more ecologically valid setting.

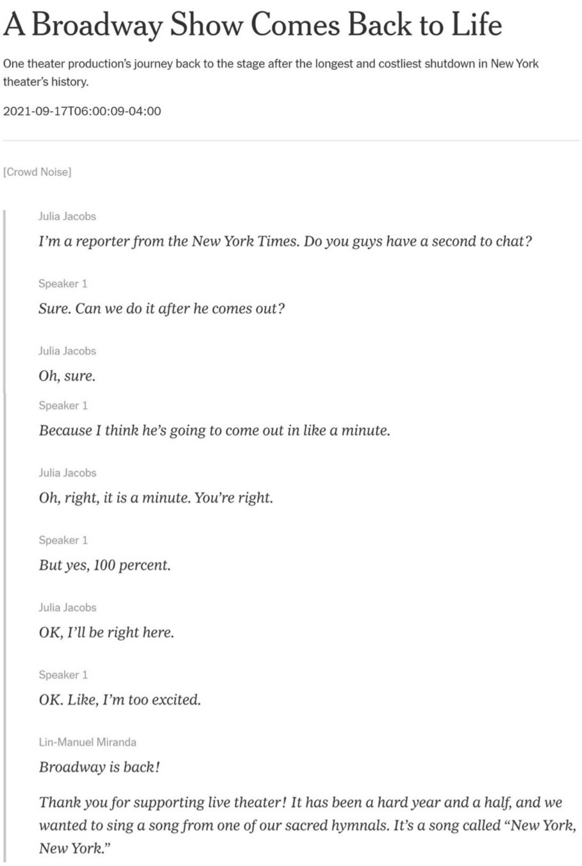
3.4. Study 3: analyzing conversations in a real podcast

We scraped and analyzed transcripts and behavioral data from *The Daily* podcast.

Sample and Data Scraping. Episodes of *The Daily* are released on weekdays (from Monday to Friday, by 6 am), with an additional Sunday special ("*The Sunday Read*"). The typical episode duration ranges from a minimum of 20 to approximately 40 minutes.

Episodes follow an interview format in which hosts interview journalists and experts, summarizing and commenting on stories that are either currently in the news or of particular interest to the audience. Conversations vary in terms of concentration: some are dominated by a few speakers, while others are fragmented, with speech distributed more evenly among multiple speakers (number of speakers: $M = 8.90$, $SD = 6.26$, min = 2, max = 56). This variability makes the podcast an ideal context for examining the effects of discourse concentration.

We scraped episode transcripts from the *New York Times*' website. As reported on the website, transcripts are generated using



A Broadway Show Comes Back to Life

One theater production's journey back to the stage after the longest and costliest shutdown in New York theater's history.

2021-09-17T06:00:09-04:00

[Crowd Noise]

Julia Jacobs

I'm a reporter from the New York Times. Do you guys have a second to chat?

Speaker 1

Sure. Can we do it after he comes out?

Julia Jacobs

Oh, sure.

Speaker 1

Because I think he's going to come out in like a minute.

Julia Jacobs

Oh, right, it is a minute. You're right.

Speaker 1

But yes, 100 percent.

Julia Jacobs

OK, I'll be right here.

Speaker 1

OK. Like, I'm too excited.

Lin-Manuel Miranda

Broadway is back!

Thank you for supporting live theater! It has been a hard year and a half, and we wanted to sing a song from one of our sacred hymnals. It's a song called "New York, New York."

Fig. 4. Excerpt of a transcript of one of *The Daily*'s episodes (retrieved from https://www.nytimes.com/2021/09/17/podcasts/the-daily/broadway-returns-six-musical.html).

speech-recognition software and reviewed by human transcribers to ensure accuracy in spelling and speakers' turns. The transcripts of the episodes provide detailed information on speakers' turns, allowing us to identify who says what. This clear subdivision of speakers' turns (see Fig. 4) enabled the precise computation of discourse concentration for each episode, based on the relative length of the speakers' interventions in terms of characters.

We collected data on the dependent variable, engagement, from the podcast hosting platform Castbox, which reports the exact number of likes and comments per episode. In October 2022, we scraped engagement-related metrics (likes and comments) for each of the 401 episodes released between January 2021 (when Castbox began collecting engagement metrics for this podcast) and September 2022. We extracted engagement-related data one month after the last scraped episode, to exhaust the hype effect and obtain more stable measures of engagement.

Measures. We measured discourse concentration using the HHI, computed based on the relative length of speakers' turns for each episode (see Study 1). We computed the length of speakers' contributions in terms of the number of characters per intervention. We measured content *valence* using the LIWC 2022 Tone variable (Boyd et al., 2022). In line with prior research, we measured overall engagement as the sum of likes and comments per episode (Kanuri et al., 2024; Weiger et al., 2025).

Control Variables. To reduce the risk of potential confounding factors, we also measured a series of control variables. Because conversational outcomes may depend on the discussed content (Packard & Berger, 2021), we controlled for the *topics* covered in the episodes. We performed a Latent Dirichlet Allocation (LDA; Blei et al., 2003) analysis of the 401 transcripts to discover common topics. After a pre-treatment phase and a subsequent refinement, a 4-topic solution emerged: Economics (representative words: money, economy, company), internal affairs (democrat, Trump, election), international issues (war, Russia, Ukraine), and health (Covid, vaccine, health). Therefore, we created three topic-dummy variables using economics as the baseline. To control for *audio features* (Hildebrand et al., 2020), we considered two variables computed at the episode level[1]: *average pitch* (F0), measured with the YIN frequency estimator, where low values describe low pitch and high values describe high pitch; and *harmonicity*, measured with the HNR (Harmonics-to-Noise Ratio) index, which indicates the degree of acoustic periodicity expressed in decibels (dB) – low values suggest equal energy in the harmonics and noise, whereas high values suggest predominant periodicity. We computed both measures using the Python Librosa package. To control for *episode and guests' features*, we considered the episode duration in seconds, airing temporal order of episodes (trend), the number of recorded turns over the total turns, the average guest speaker popularity and the dominant guest speaker popularity (operationalized based on the character length of the guests' Wikipedia pages – Laouenan et al., 2022), and the ratio of guest contributions to host contributions. Additionally, we measured the *temporal* perspective of the episode using the LIWC 2022 past and future variables. Finally, we took into consideration *psycho-linguistic variables* examined in previous studies on user engagement: *authenticity* (Moon et al., 2021), episode's average *concreteness* and *familiarity* (Paetzold & Specia, 2016), *arousal* (Mohammad, 2018), the *variability of valence* along the episode (the standard deviation of the valence scores over the episode's turns), and *language style matching* (LSM – Boyd et al., 2022). To facilitate interpretation, we mean–centered the independent, moderating, and control variables.

Results. Descriptive statistics and correlations for all the variables are reported in Table 2. The results of regression analyses are reported in Table 3.

In both interaction models, without ($b_{interaction} = 0.51$, $p = .0008$) or including ($b_{interaction} = 0.58$, $p < .001$) control variables, we observed a significant and positive discourse concentration × valence interaction. Results also showed a positive effect of episode duration and an increasing trend of engagement across episodes. We probed the interaction effect with a Johnson-Neyman analysis on the complete interaction model. Results showed that for more negatively-valenced episodes (up to the 25th percentile of valence) the effect of discourse concentration on engagement was negative. In contrast, for more positively-valenced episodes (from the 85th percentile of valence) the effect of discourse concentration on engagement was positive. We computed the marginal effects of simple slopes (at M ± 1SD of variables) and found that a 10% increase in discourse concentration produces a 3.09% increase in engagement for positively-valenced episodes, but a 3.95% decrease for negatively-valenced episodes. Fig. 5 illustrates the moderating role of content valence on the effect of discourse concentration on engagement. These results support H1.

Checks of Assumptions and Robustness Analyses. We conducted a set of additional analyses to ensure that OLS estimates were valid. First, the highest VIF was 2.52, suggesting non-problematic multicollinearity. This conclusion is also supported by the stability of the focal interaction effect across models with or without control variables. Second, neither the complete linear model ($\rho = 0.03$, $p = .51$) nor the complete interaction model ($\rho = 0.01$, $p = .92$) suffer from autocorrelation. Third, Breusch-Pagan tests of heteroskedasticity on the complete linear model (χ^2 (21) = 24.26, p = .28) and on the complete interaction model (χ^2 (22) = 26.22, p = .24) suggested homoscedastic errors, supporting the confidence in the OLS standard errors and inferential conclusions. Fourth, considering that the distributions of the three core variables show some potential outliers (engagement has six standardized values >2.58; discourse concentration has six standardized values < -2.58; valence has four standardized values >2.58), we performed additional regression analyses. Using logarithmic transformations ($b = 0.06$, $p < .001$), 90% winsorized data ($b = 0.73$, $p < .001$), and 95% winsorized data ($b = 0.71$, $p < .001$) the interaction effect remained positive and significant. Fifth, we checked for any potential endogeneity issues. First, we used the Gaussian copula approach (Becker et al., 2022) and created copula terms for discourse concentration and valence. A regression model including these two terms still showed a significant and positive interaction term ($b = 0.59$, $p = .001$) and non-significant copula terms. This evidence holds even when using bootstrap standard errors. We acknowledge that the non-normal dependent variable (and therefore non-normal residuals) may bias the copula approach. Therefore, we applied the control function

[1] Currently no reliable and open technologies are available to obtain precise measures of audio features per speaker when a conversation includes several speakers.

Table 2
Descriptive statistics and correlations.

	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Engagement	9.15	5.05	1.00																					
2. Discourse concentration	0.45	0.12	−0.02	1.00																				
3. Valence	25.53	12.38	−0.10	−0.17	1.00																			
4. Internal affairs	0.30	0.46	−0.09	−0.16	−0.04	1.00																		
5. International issues	0.26	0.44	0.03	0.25	−0.29	−0.38	1.00																	
6. Health	0.25	0.43	0.03	−0.19	0.13	−0.38	−0.34	1.00																
7. YIN	541.00	13.71	−0.13	−0.04	0.05	0.03	−0.28	0.19	1.00															
8. HNR	1.43	1.35	0.00	0.13	−0.06	−0.19	−0.01	0.24	0.17	1.00														
9. Duration (in *sec.*)	1879.99	378.93	0.18	−0.32	0.14	0.03	−0.03	0.03	−0.06	−0.05	1.00													
10. Trend (t)	201.00	115.90	0.31	0.08	−0.05	−0.08	0.09	−0.01	−0.28	−0.03	0.07	1.00												
11. Recorded turns/total turns	0.25	0.17	0.04	−0.61	0.15	0.15	−0.15	0.07	−0.02	−0.23	0.27	−0.05	1.00											
12. Av. guest popularity	3896.50	3182.73	−0.01	0.05	0.11	0.17	0.01	−0.21	−0.15	−0.05	0.00	−0.05	0.00	1.00										
13. Dominant sp. popularity	798.32	2370.46	0.06	−0.03	0.07	0.14	−0.04	−0.04	−0.07	−0.09	0.05	−0.02	0.11	0.31	1.00									
14. Guests/Hosts contributions	4.02	3.92	0.01	−0.33	0.17	−0.03	−0.05	0.13	0.05	−0.04	0.34	−0.01	0.36	−0.12	−0.03	1.00								
15. Focus past	4.69	1.50	0.03	−0.18	−0.15	−0.05	0.18	0.07	0.01	0.01	0.27	−0.01	0.22	−0.10	−0.14	0.20	1.00							
16. Focus future	1.64	0.55	−0.04	−0.03	0.12	−0.04	−0.17	0.11	0.03	0.12	−0.08	−0.05	−0.17	−0.03	0.09	−0.04	−0.37	1.00						
17. Authentic	4.62	16.01	0.03	−0.19	0.07	−0.23	−0.01	0.27	0.02	0.04	0.10	−0.01	0.05	−0.20	0.03	0.18	0.03	0.28	1.00					
18. Concreteness	324.51	5.44	0.03	−0.27	0.04	−0.07	0.05	0.03	−0.07	−0.22	0.22	0.03	0.46	−0.10	−0.10	0.23	0.36	−0.30	0.03	1.00				
19. Familiarity	583.76	5.20	0.11	−0.18	0.13	−0.38	−0.15	0.41	0.12	0.12	0.26	0.03	0.11	−0.27	−0.02	0.22	0.26	0.17	0.45	0.20	1.00			
20. Arousal	0.49	0.01	0.04	0.15	−0.53	0.28	0.32	−0.35	−0.17	−0.07	−0.11	0.08	−0.08	0.11	0.03	−0.16	0.08	−0.20	−0.39	−0.08	−0.47	1.00		
21. Variability of valence	3.20	3.34	−0.07	−0.20	0.53	0.04	−0.09	0.05	−0.03	−0.04	0.12	−0.04	0.17	0.16	0.09	0.13	−0.15	0.09	−0.05	−0.02	−0.01	−0.13	1.00	
22. LSM	0.72	0.11	−0.11	0.14	0.00	−0.01	−0.08	0.03	0.11	0.09	−0.14	−0.18	−0.31	−0.01	−0.06	−0.10	−0.13	0.03	−0.10	−0.25	−0.06	0.00	−0.01	1.00

Correlations larger than |0.128| are significant at the $p < .01$ level. Correlations larger than |0.098| are significant at the $p < .05$ level.

Table 3
Results of regression analyses.

	Linear model, only main variables			Interaction model, only main variables			Linear model, all variables			Interaction model, all variables		
	b		*s.e.*	*b*		*s.e.*	*b*		*s.e.*	*b*		*s.e.*
Intercept	9.15	***	0.25	9.28	**	0.25	9.15	***	0.24	9.29	***	0.24
Main Variables												
Discourse concentration	-1.41		2.11	- 2.02		2.09	−0.49		2.73	−0.88		2.68
Valence	- 0.04	*	0.02	- 0.03		0.02	−0.04		0.03	−0.02		0.03
Disc. concentration x Valence				0.51	***	0.15				0.58	***	0.15
Control variables												
Topics												
Topic: internal affairs							−1.33		0.82	−1.09		0.81
Topic: international issues							−0.92		0.86	−0.40		0.86
Topic: health							−0.32		0.77	0.04		0.76
Audio features												
YIN							−0.02		0.02	−0.01		0.02
HNR							0.04		0.19	0.04		0.19
Episode and guests' features												
Duration (in sec.)							0.00	**	0.00	0.00	**	0.00
Trend							0.01	***	0.00	0.01	***	0.00
Recorded turns/total turns							1.56		2.08	0.99		2.05
Average guest popularity							0.00		0.00	0.00		0.00
Dominant speaker popularity							0.00		0.00	0.00		0.00
Guests/Hosts contributions							−0.05		0.07	0.01		0.07
Time perspective												
Focus past							−0.19		0.20	−0.20		0.20
Focus future							−0.39		0.52	−0.47		0.51
Psycholinguistics												
Authentic							−0.00		0.02	0.01		0.02
Concreteness							−0.04		0.06	−0.03		0.05
Familiarity							0.10		0.07	0.10		0.07
Arousal							55.78		68.54	75.74		67.55
Variability of valence							−0.07		0.09	−0.06		0.09
LSM							−1.56		2.32	−1.92		2.28
R^2		0.01			0.04			0.16			0.19	
F		2.12			5.24**			3.47**			4.09**	
ΔR^2		–			0.03**			–			0.03**	

Note: * $p < .05$, ** $p < .01$, *** $p < .001$. All predictors are mean-centered.

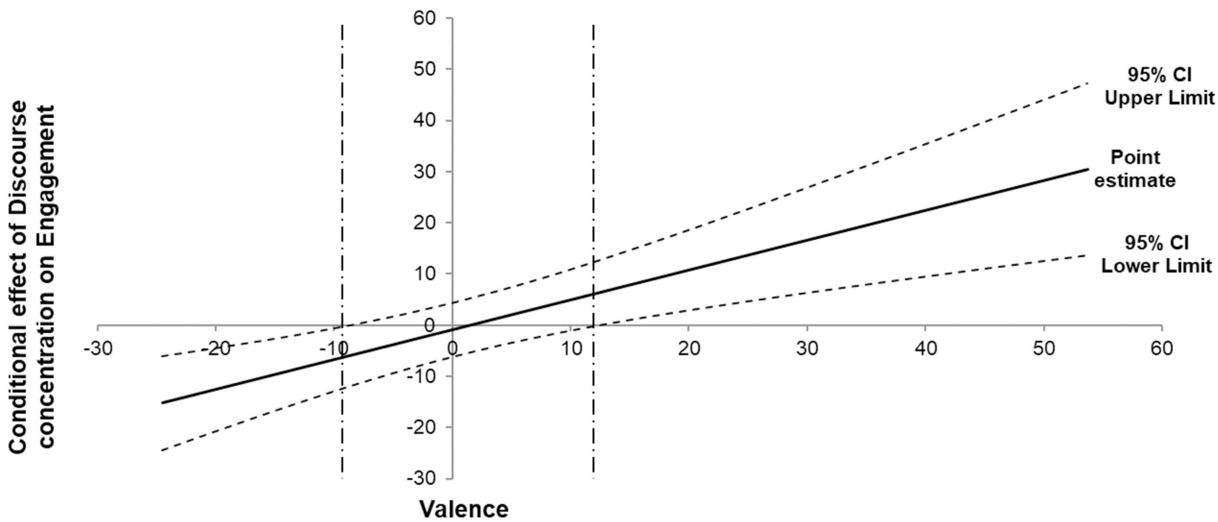


Fig. 5. Effect of discourse concentration on engagement conditional on valence (Study 3).

approach (Sande & Ghosh, 2018) estimating first-stage equations for discourse concentration and valence as functions of the control variables (results are consistent also without using control variables), the respective lagged variables, the number of speakers, and average TF × IDF (Robertson, 2004), a popular measure of word relevance. Both first-stage equations show good R-squared values

(0.55 and 0.53, respectively) with VIFs lower than 2.82, suggesting adequate coverage of the key variables. In the second-stage models, we included the first-stage equations' residuals (which were uncorrelated, $r = 0.03$). Results again show a significant and positive interaction term ($b = 0.53$, $p = .001$). Overall, these results suggest that the model does not suffer from endogeneity.

To check the robustness of the results, we re-estimated the complete interaction model using alternative measures of discourse concentration and valence. We computed two alternative measures of discourse concentration (beyond HHI): a conversation dominance index, using the formula introduced in Study 1; and an entropy index, computed as $-\sum_{s=1}^{S} p_s \times \ln(p_s)$. Higher values of the entropy index imply higher fragmentation, while lower values indicate higher concentration. The three discourse concentration measures are strongly and significantly correlated with each other ($|r| = [0.85, 0.96]$), demonstrating concurrent validity. As alternative measures of valence (beyond the LIWC 2022 Tone variable), we used the Evaluative Lexicon Suite valence measure (Rocklage et al., 2018), a percentage difference score between positive and negative turns predicted by the RoBERTa sentiment analysis tool (Liu et al., 2019), and the valence measure from the NRC VAD Lexicon dictionary (Mohammad, 2018). The four measures of content valence show medium-to-high, significant correlations ($r = [0.41, 0.75]$), demonstrating concurrent validity.

Across all potential measures, the focal interaction remains consistently significant and in the expected direction, both with and without control variables. Table 4 synthesizes the robustness analyses for the focal interaction in the complete models. These results confirm the robustness of the presented evidence supporting H1.[2]

Finally, we measured engagement as the sum of likes and comments, consistent with prior research (e.g., Kanuri et al., 2024). However, since comments can vary in valence, one might question their inclusion in this metric. To address this concern, we re-estimated the full interaction model using only likes as the measure of engagement and found that the positive and significant interaction effect remained stable ($b = 0.53$, $p < .001$).

Discussion. Study 3 demonstrates that the effect of discourse concentration on engagement is moderated by valence in a more ecologically valid setting. When an episode is primarily positively-valenced, the audience engages more with a concentrated conversation, with one or a few speakers dominating the conversation and owning a larger share of the spoken contributions. Conversely, when an episode is primarily negatively-valenced, the audience engages more with a fragmented conversation, with speech distributed more evenly across speakers.

4. General discussion

Conversational content has become a prominent form of digital communication, with podcasts representing its most paradigmatic form. Despite a rich literature on conversation analysis (Grice, 1975; Hutchby, 2019), marketing research on what makes conversational content engaging remains limited. Podcasts, despite their mainstream reach and economic relevance, are no exception.

Building upon conversation analysis studies, we close this gap by developing and testing a conceptual model showing that discourse concentration (concentrated vs. fragmented conversation) and content valence (positive or negative) jointly influence audience engagement, mediated by empathy toward the speakers. Across three complementary studies, using transcript-based stimuli, AI-generated audio stimuli, and the econometric analysis of 401 real podcast episodes, we find general support for our conceptual model, showing that when the content conveyed in a podcast episode is positively-valenced, the audience is more engaged when the conversation has one or a few speakers dominating the conversation. Conversely, when the content conveyed is negatively-valenced, the audience is more engaged with a more fragmented conversation, with speakers sharing spoken contributions more evenly. Study 1 and Study 2 provide causal evidence for this conditional effect in controlled settings using text-based and AI-generated audio stimuli, respectively, and document the mediating role of empathy toward the speakers. Study 3 features the econometric analysis of transcripts from episodes of a real podcast, thereby providing ecological validity for the proposed moderation effect. Remarkably, the econometric model explains nearly 20% of the variance in audience engagement. This result underscores the substantive importance of conversational structure and its interaction with content valence in explaining engagement over and above content characteristics and speaker-related factors.

4.1. Theoretical contributions

We offer three contributions to the literature on conversation analysis and language research in marketing contexts. First, we respond to calls for further research on interpersonal communication settings by extending language-based consumer behavior frameworks to multi-speaker conversational contexts. In their bibliographic analysis on the emergence and evolution of consumer language research, Packard and Berger (2024) highlight the need to study language in emerging channels (i.e., beyond reviews and social media) and in spoken interactions. In a related vein, research on unstructured data emphasizes the importance of understanding speaker interaction dynamics and their impact on behavioral outcomes (Grewal et al., 2022). In response to these calls, we introduce discourse concentration, a core construct capturing how conversational floor time is allocated across speakers. Discourse concentration has theoretical roots in adjacent disciplines (Mast, 2002; Reid & Ng, 2000), but it has remained largely unexplored as a driver of audience engagement in marketing contexts.

[2] Since previous research has found a numerosity effect in marketing videos (Chang et al., 2023), we estimated a model including the number of speakers as a control variable. The focal interaction effect remained significant ($b = 0.55$, $p < .001$). The number of speakers had only a marginally significant effect on engagement ($b = -0.12$, $p = .07$). As expected, the number of speakers is strongly correlated with discourse concentration ($r = -0.57$, $p < .001$); therefore, we preferred to exclude it from the main analyses to avoid multicollinearity.

Table 4
Robustness analyses – the Discourse Concentration × Content Valence interaction effect using different measures.

		Content Valence measures			
		Tone LIWC 2022	Valence (Rocklage et al., 2018)	RoBERTa valence measure	Valence (Mohammad, 2018)
Discourse Concentration measures	**HHI**	$b = 0.58$, $p = .000$	$b = 9.90$, $p = .001$	$b = 38.82$, $p = .006$	$b = 902.74$, $p = .002$
	Dominance index	$b = 0.55$, $p = .000$	$b = 9.54$, $p = .001$	$b = 38.42$, $p = .004$	$b = 928.01$, $p = .001$
	Entropy	$b = -0.17$, $p = .001$	$b = -2.83$, $p = .003$	$b = -11.29$, $p = .011$	$b = -215.87$, $p = .016$

Prior research has begun to explore the role of multiple voices in persuasion. Chang et al. (2023) found that messages featuring more (vs. fewer) voices are more persuasive in Kickstarter campaigns and video advertising. We extend this stream of research by analyzing the effects of discourse concentration, a construct that moves beyond the nominal count of voices to capture the distribution of speakers' participation (concentrated vs. fragmented). Importantly, we demonstrate that the effect of discourse concentration depends on content valence. Drawing on insights from conversation analysis concerning asymmetries in speaker participation (Hutchby, 1996) and research on communication and persuasion (e.g., Ratneshwar & Chaiken, 1991), this research shows how audiences jointly process conversational structure and content valence when engaging with conversational content. By demonstrating that the same conversational structure can either increase or decrease engagement depending on content valence, our findings highlight the importance of considering structure–content fit. To the best of our knowledge, this research represents one of the first systematic efforts to integrate concepts from conversation analysis into marketing research. In doing so, we open a dialogue between conversation analysis and marketing research.

Second, we identify empathy toward the speakers as the psychological mechanism explaining why the effect of discourse concentration on engagement depends on content valence (Hoffman, 1984). In doing so, our work extends research on empathy in marketing contexts (Bagozzi & Moore, 1994; Gelbrich et al., 2021) by showing that empathic responses are shaped not only by message content but also by the structural properties of the conversation through which that content is delivered. These findings extend current theories on how affective and structural features of conversational content jointly shape audience responses (Collins & Miller, 1994; Lieberman et al., 2022).

Third, from a substantive perspective, we contribute to marketing research on podcasts as cultural products (Ghasemi et al., 2026), which remains surprisingly limited despite the format's rapid growth and economic relevance. By operationalizing discourse concentration and content valence at scale across 401 real podcast episodes, we demonstrate that constructs from conversation analysis can be measured and tested in naturalistic media environments. The full model including the interaction between conversational structure and content valence, and a rich set of content, audio, and speaker characteristics, explains nearly 20% of the variance in real-world audience engagement. This result establishes that the structural properties of conversation have meaningful and measurable consequences for audience behavior beyond what controlled settings alone can reveal. Our findings suggest that podcast effectiveness is not determined solely by topic or speaker characteristics, but also by how conversational floor time is distributed across participants. In doing so, we stimulate research on podcasts grounded in conversational constructs.

4.2. Managerial contribution

The results of this research offer actionable insights for conversational content platforms, producers, and creators. When designing podcast episodes, livestreamed discussions on social media, or roundtable-style programs, producers and creators should consider the predominant valence of the content to be covered when deciding the optimal distribution of conversational floor among speakers.

For conversational content with more positive themes, such as those focused on self-improvement, our research suggests that a concentrated conversational structure may be more effective. This approach, featuring fewer dominant speakers, fosters stronger empathic bonds between the audience and the speakers due to its more intimate nature, which in turns enhances engagement. Conversely, conversational content dealing with more negative or challenging topics, such as true crime stories, may benefit from a more balanced conversational structure involving several speakers. This approach reduces the risk of "shooting the messenger" mechanisms (John et al., 2019) and activates social validation and support, which help audiences empathize with the speakers when confronting difficult content, making it easier to process and engage with the content. Notably, the field study results suggest that managing discourse concentration may be particularly consequential for negatively-valenced content: a 10% increase in discourse concentration produces a 3.09% increase in engagement for positive episodes, but a 3.95% decrease for negative episodes. This pattern, consistent across our experimental studies, suggests that podcast producers and content creators dealing with negative topics stand to benefit most from actively distributing conversational floor time across multiple speakers.

To illustrate these findings, we can examine successful conversational content. On the one hand, "*New Heights with Jason and Travis Kelce*" is a show with mostly positively-valenced episodes that thrives on a concentrated format. The podcast capitalizes on the natural chemistry and humor of the two NFL stars and podcast hosts, creating an engaging discussion on current sports news. The dominant presence of the Kelce brothers allows the audience to form a strong connection with their personalities and perspectives. On the other hand, "*Dateline NBC*" represents a negatively-valenced podcast that benefits from a more balanced approach, being an acclaimed true crime show derived from a longstanding television series that delivers investigations into real-life crime cases. Each episode typically features interviews with key participants, including law enforcement officials, family members, and occasionally the victims themselves. The cases are meticulously researched and presented in a documentary-like format that maintains the audience's interest. With leading correspondents Keith Morrison, Andrea Canning, and Josh Mankiewicz, the podcast upholds the high standards of investigative journalism that audiences have come to expect from the television show, benefiting from the numerous perspectives brought to the table by the many speakers involved.

Our findings also offer implications for companies seeking to sponsor or advertise on conversational content. Highly engaged podcast audiences tend to pay greater attention to embedded advertisements and exhibit stronger advertising receptivity (Brinson & Lemon, 2023). Beyond considering the overall popularity of podcasts and livestreams, advertisers should take into account both the conversational structure and the content valence of individual episodes. Specifically, brands may benefit from sponsoring more concentrated conversational formats when episodes focus on positively-valenced topics, such as self-help, personal improvement, or comedy. Conversely, for negatively-valenced episodes addressing topics like crime, war, political crises, or health emergencies, advertisers may prefer more fragmented conversational structures, which appear to foster greater engagement.

4.3. Limitations and directions for future research

Like any research, this study has some limitations that open avenues for future work. First, our findings require further generalization. While we focus on podcasts as a representative form of conversational content, future research could examine whether and how the observed effects generalize to other conversational media formats (Grewal et al., 2022). Differences in visual cues in videos related to speakers' aesthetics (e.g., DeShields Jr et al., 1996) may introduce important boundary conditions for the effects of conversational structure on engagement.

Second, while the present research focuses on discourse concentration, future work could examine how other structural features of conversation influence audience engagement. Prior research has highlighted the roles of turn-taking dynamics and sequencing, speaker order and speaking time, and conversational interactivity in shaping conversational experiences (e.g., Parker, 1988; Rühlemann & Gries, 2015). Investigating how these features interact with content characteristics such as valence, uncertainty, or controversy could further advance understanding of conversational engagement. The study of conversational structure requires analyzing spoken contributions with clear separation of speakers' turns. Relatedly, future research could examine whether the effects of discourse concentration vary as a function of baseline group size. The range of possible HHI values is mathematically constrained by the number of speakers, and in our field data the number of speakers is strongly correlated with discourse concentration, making the two variables difficult to disentangle. Our experimental studies held the number of speakers constant at six, which is close to the field study median, but future experimental work could directly manipulate group size to examine whether it moderates the effects we document.

Third, our research focuses on conversational content, which, by definition, involves two or more speakers. Prior research, however, has distinguished between monologues and dialogues as qualitatively different communication formats (e.g., Fay et al., 2000). Future research could further investigate the conditions under which monologic versus dialogic structures are more persuasive. For example, research on the identifiable victim effect suggests that people often respond more strongly to a single, identifiable victim than to statistical or abstract victims, with single, vivid narratives eliciting greater empathy, compassion, and willingness to help (Kogut & Kogut, 2013; Small & Loewenstein, 2003). Importantly, the psychological mechanisms emphasized in the identifiable victim literature may operate differently in conversational settings, where multiple speakers, social validation, and interactional dynamics shape how negative information is processed. Future research could therefore investigate the boundary conditions under which concentrated versus fragmented conversational structures are more effective in negatively-valenced contexts, including comparisons between monologic and dialogic formats.

Fourth, while in the field study we controlled for episode-level audio features (pitch and harmonicity), we could not analyze these features at the speaker-level as reliable speaker diarization technology was not available for episodes with large numbers of speakers at the time of data collection. Advances in machine learning and speech processing may soon enable more granular analyses of audio features and their effects on engagement.

Fifth, while our field study controls for speaker characteristics including average and dominant guest popularity and the ratio of guest to host contributions, we acknowledge that speaker-based heterogeneity represents a broader challenge that our design cannot fully address. Specifically, the effects of discourse concentration may vary across different speaker roles or identities in ways that our proxies do not fully capture. To explore this possibility, we conducted additional robustness analyses including three-way interaction models testing whether the discourse concentration × content valence interaction is moderated by the ratio of recorded to total turns ($b = -0.11, p = .88$), average guest speaker popularity ($b = -0.00, p = .45$), dominant guest speaker popularity ($b = 0.00, p = .42$), and the ratio of guest to host contributions ($b = -0.04, p = .35$). None of these interactions reached statistical significance, suggesting that the core interaction effect is robust to variation in speaker prominence and role composition. Full results are reported in Web Appendix 5. We acknowledge, however, that these proxies capture speaker prominence rather than speaker-specific roles, and that the interpretation of conversational dominance may differ depending on the role a speaker occupies and the expectations attached to it. For instance, a single expert dominating a conversation may be perceived differently than a host dominating the same exchange, representing a plausible boundary condition on our effects that our current design cannot test. Future research could address this limitation by directly manipulating speaker roles across different conversational contexts to examine whether role-based expectations moderate the effects of discourse concentration on engagement.

Sixth, we propose perceived empathy toward the speakers as the main mediator of the analyzed interaction effect. Empathy toward the speakers is a theoretically meaningful mediator that helps explain the two simple effects of discourse concentration observed for positively- and negatively-valenced content. However, we acknowledge that it may operate in conjunction with other processes, such as *processing fluency* (Reber et al., 2004). For example, empathizing with speakers may facilitate more fluent elaboration of conversational content, which could in turn increase engagement. Future research may investigate further the mechanisms underlying the interaction between discourse concentration and content valence, testing for more complex, serial mediation models.

Finally, this research focuses on content valence as the primary moderator. Future research may consider other potential moderators of the effect of discourse concentration, such as the type of content or individual differences. Exploratory analyses on Study 3 data suggest that discourse concentration and content valence positively interact with the international issues versus economy topic dummy ($b = 1.38$, $p = .006$). This evidence suggests that the core interaction may be stronger for specific topics. Additionally, individual differences may moderate the core interaction effect. For example, individuals higher in need for cognition may be better equipped to process information from multiple speakers, and therefore show stronger engagement with fragmented structures regardless of valence compared to individuals lower in need for cognition. More broadly, individual differences in empathic concern or emotion regulation may influence how strongly conversational structure influences engagement with negatively-valenced content in particular. Future research could examine these moderators to identify the boundary conditions of the proposed effects.

Despite these limitations, this research advances our understanding of how conversational structure influences audience engagement, with implications for theory, practice, and the growing body of marketing research on conversational media.

CRediT authorship contribution statement

Gaetano Miceli: Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Ludovica Serafini:** Writing – original draft, Software, Methodology, Investigation, Data curation. **Ernesto Cardamone:** Writing – review & editing, Software, Methodology, Data curation. **Maria Antonietta Raimondo:** Writing – review & editing, Conceptualization. **Francisco Villarroel Ordenes:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Irene Scopelliti:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to thank Shankha Basu and two anonymous IJRM reviewers, the Marketing Group at Vrije Universiteit Amsterdam, attendees of the "Marketing and the Creator Economy" Conference held at Columbia Business School, and participants in the European Quant Marketing Workshop for their insightful feedback on earlier versions of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2026.07.005.

Data availability

Data will be made available on request.

References

Alantari, H. J., Currim, I. S., Deng, Y., & Singh, S. (2022). An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *International Journal of Research in Marketing, 39*(1), 1–19.

Ames, K. (2013). Two hosts and a caller: Analysing call sequences in a dual-host radio talkback setting. *Discourse Studies, 15*(3), 263–277.

Asher, N., & Lascarides, A. (2013). Strategic conversation. *Semantics and Pragmatics, 6*, 1–62. Article 2.

Bagozzi, R. P., & Moore, D. J. (1994). Public service advertisements: Emotions and empathy guide prosocial behavior. *Journal of Marketing, 58*(1), 56–70.

Barasch, A., & Berger, J. (2014). Broadcasting and narrowcasting: How audience size affects what people share. *Journal of Marketing Research, 51*(3), 286–299.

Bastos, W. (2020a). Want to make me happy? Tell me about your experiences but not your objects. *Journal of Consumer Affairs, 54*(3), 978–1001.

Bastos, W. (2020b). "Speaking of purchases": How conversational potential determines consumers' willingness to exert effort for experiential versus material purchases. *Journal of Interactive Marketing, 50*(1), 1–16.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323–370.

Becker, J. M., Proksch, D., & Ringle, C. M. (2022). Revisiting Gaussian copulas to handle endogenous regressors. *Journal of the Academy of Marketing Science, 50*(1), 46–66.

Behrens, R., Kupfer, A. K., & Hennig-Thurau, T. (2025). There is business like show business! What marketing scholars and managers can learn from 40 years of entertainment science research. *Journal of the Academy of Marketing Science, 53*, 760–780.

Berger, J., Kim, Y. D., & Meyer, R. (2021). What makes content engaging? How emotional dynamics shape success. *Journal of Consumer Research, 48*(2), 235–250.

Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research, 49*(2), 192–205.
Berry, R. (2020). There are just 3 types of podcast. https://richardberry.eu/there-are-just-3-types-of-podcast/. last accessed June 5, 2024.
Biber, D., Egbert, J., Keller, D., & Wizner, S. (2021). Towards a taxonomy of conversational discourse types: An empirical corpus-based analysis. *Journal of Pragmatics, 171*, 20–35.
Blanchard, S. J., Goldenberg, J., Pauwels, K., & Schweidel, D. A. (2022). Promoting data richness in consumer research: How to develop and evaluate articles with multiple data sources. *Journal of Consumer Research, 49*(2), 359–372.
Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(Jan), 993–1022.
Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The development and psychometric properties of LIWC-22* (pp. 1–47). Austin, TX: University of Texas at Austin.
Brinson, N. H., & Lemon, L. L. (2023). Investigating the effects of host trust, credibility, and authenticity in podcast advertising. *Journal of Marketing Communications, 29*(6), 558–576.
Cascio Rizzo, G. L., Berger, J., & Zhou, M. (2026). Talking with your hands: How hand gestures influence communication. *Journal of Marketing Research, 63*(3), 518–537.
Cascio Rizzo, G. L., Villarroel Ordenes, F., Pozharliev, R., De Angelis, M., & Costabile, M. (2024). How high-arousal language shapes micro-versus macro-influencers' impact. *Journal of Marketing, 88*(4), 107–128.
Chang, H. H., Mukherjee, A., & Chattopadhyay, A. (2023). More voices persuade: The attentional benefits of voice numerosity. *Journal of Marketing Research, 60*(4), 687–706.
Chaudhry, A., Wang, Y., & Ouyang, E. (2025). Community gifting and recipients' engagement with peers and influencers: Evidence from natural experiments on Twitch. *Journal of Interactive Marketing, 60*(1), 83–99.
Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin, 116*(3), 457–475.
Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 71–100.
DeShields, O. W., Jr., Kara, A., & Kaynak, E. (1996). Source effects in purchase decisions: The impact of physical attractiveness and accent of salesperson. *International Journal of Research in Marketing, 13*(1), 89–101.
Efthymiou, F., Hildebrand, C., de Bellis, E., & Hampton, W. H. (2024). The power of AI-generated voices: How digital vocal tract length shapes product congruency and ad performance. *Journal of Interactive Marketing, 59*(2), 117–134.
Fay, N., Garrod, S., & Carletta, J. (2000). Group discussion as interactive dialogue or as serial monologue: The influence of group size. *Psychological Science, 11*(6), 481–486.
Gelbrich, K., Hagel, J., & Orsingher, C. (2021). Emotional support from a digital assistant in technology-mediated services: Effects on customer satisfaction and behavioral persistence. *International Journal of Research in Marketing, 38*(1), 176–193.
Ghasemi, H., Astvansh, V., & Sephri, A. (2026). How should content creators narrate their content? The impact of emotionality flips on audience engagement. *Journal of the Academy of Marketing Science*. forthcoming.
Gnewuch, U., Morana, S., & Maedche, A. (2017). *Towards designing cooperative and social conversational agents for customer service. ICIS* (pp. 1–13).
Gonzàlez-Lloret, M. (2010). Conversation analysis and speech act performance. In *Speech Act Performance: Theoretical, Empirical and Methodological Issues* (pp. 57–73).
Grewal, D., Herhausen, D., Ludwig, S., & Villarroel Ordenes, F. (2022). The future of digital communication research: Considering dynamics and multimodality. *Journal of Retailing, 98*(2), 224–240.
Grewal, D., Ludwig, S., Herhausen, D., Grewal, L., Ordenes, F. V., Kopalle, P. K., & Bohling, T. (2026). Express: How frontline employees' relational communication in online service interactions drives customer satisfaction. *Production and Operations Management, 35*(8), 2963–2980.
Grice, H. P. (1975). Logic and conversation. In *Speech acts* (pp. 41–58). Brill.
Hartmann, J., Huppertz, J., Schamp, C., & Heitmann, M. (2019). Comparing automated text classification methods. *International Journal of Research in Marketing, 36*(1), 20–38.
Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). New York: Guilford Press.
Henderson, P. W., & Cote, J. A. (1998). Guidelines for selecting or modifying logos. *Journal of Marketing, 62*(2), 14–30.
Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W. H., Hoffman, D. L., & Novak, T. P. (2020). Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications. *Journal of Business Research, 121*, 364–374.
Hoffman, M. L. (1984). Interaction of affect and cognition in empathy. In C. Izard, J. Kagan, & R. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 103–131). New York: Cambridge University Press.
Holler, J., Alday, P. M., Decuyper, C., Geiger, M., Kendrick, K. H., & Meyer, A. S. (2021). Competition reduces response times in multiparty conversation. *Frontiers in Psychology, 12*, Article 693124.
Humphreys, A., & Wang, R. J. H. (2018). Automated text analysis for consumer research. *Journal of Consumer Research, 44*(6), 1274–1306.
Hutchby, I. (1996). *Confrontation talk: Arguments, asymmetries and power on talk radio*. Mahwah NJ: Lawrence Erlbaum Associates.
Hutchby, I. (2019). Conversation analysis. In P. Atkinson, S. Delamont, A. Cernat, et al. (Eds.), *SAGE research methods foundations* (pp. 1–21). London: SAGE Publications.
John, L. K., Blunden, H., & Liu, H. (2019). Shooting the messenger. *Journal of Experimental Psychology: General, 148*(4), 644–666.
Kanuri, V. K., Hughes, C., & Hodges, B. T. (2024). Standing out from the crowd: When and why color complexity in social media images increases user engagement. *International Journal of Research in Marketing, 41*(2), 174–193.
Karagür, Z., Becker, J. M., Klein, K., & Edeling, A. (2022). How, why, and when disclosure type matters for influencer marketing. *International Journal of Research in Marketing, 39*(2), 313–335.
Kerbrat-Orecchioni, C. (2004). Introducing polylogue. *Journal of Pragmatics, 36*(1), 1–24.
Kogut, T., & Kogut, E. (2013). Exploring the relationship between adult attachment style and the identifiable victim effect in helping behavior. *Journal of Experimental Social Psychology, 49*(4), 651–660.
Laouenan, M., Bhargava, P., Eyméoud, J. B., Gergaud, O., Plique, G., & Wasmer, E. (2022). A cross-verified database of notable people. *Scientific Data, 9*(1), 290–309.
Lieberman, A., Schroeder, J., & Amir, O. (2022). A voice inside my head: The psychological and behavioral consequences of auditory technologies. *Organizational Behavior and Human Decision Processes, 170*, Article 104133.
Lindgren, M. (2014). "This Australian life": The Americanisation of radio storytelling in Australia. *Australian Journalism Review, 36*(2), 63–75.
Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *Roberta: A robustly optimized bert pretraining approach*. arXiv preprint arXiv:1907.11692.
Llinares, D. (2018). Podcasting as liminal praxis: Aural mediation, sound writing and identity. In *Podcasting: New aural cultures and digital media* (pp. 123–145).
Luminet, O., IV, Bouts, P., Delie, F., Manstead, A. S., & Rimé, B. (2000). Social sharing of emotion following exposure to a negatively valenced situation. *Cognition & Emotion, 14*(5), 661–688.
Marinova, D., de Ruyter, K., Huang, M. H., Meuter, M. L., & Challagalla, G. (2017). Getting smart: Learning from technology-empowered frontline interactions. *Journal of Service Research, 20*(1), 29–42.
Marx, J., Mirbabaie, M., Brendel, A. B., & Zander, K. (2021). *The role of parasocial interactions for podcast backchannel response*. . In AMCIS 2021 proceedings of the twenty-seventh Americas conference on information systems, Montreal.
Mast, M. S. (2002). Dominance as expressed and inferred through speaking time: A meta-analysis. *Human Communication Research, 28*(3), 420–450.
McAllister, J., Potts, A., Mason, K., & Marchant, G. (1994). Word duration in monologue and dialogue speech. *Language and Speech, 37*(4), 393–405.
Mohammad, S. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In *Proceedings of the 56th annual meeting of the Association for Computational Linguistics (volume 1: Long papers)* (pp. 174–184). ACL.
Moon, S., Kim, M. Y., & Iacobucci, D. (2021). Content analysis of fake consumer reviews by survey-based text categorization. *International Journal of Research in Marketing, 38*(2), 343–364.

Nematzadeh, A., Ciampaglia, G. L., Ahn, Y. Y., & Flammini, A. (2019). Information overload in group communication: From conversation to cacophony in the twitch chat. *Royal Society Open Science, 6*(10), 1–14.

Overgoor, G., Rand, W., van Dolen, W., & Mazloom, M. (2022). Simplicity is not key: Understanding firm-generated social media images and consumer liking. *International Journal of Research in Marketing, 39*(3), 639–655.

Packard, G., & Berger, J. (2021). How concrete language shapes customer satisfaction. *Journal of Consumer Research, 47*(5), 787–806.

Packard, G., & Berger, J. (2024). The emergence and evolution of consumer language research. *Journal of Consumer Research, 51*(1), 42–51.

Paetzold, G., & Specia, L. (2016). Inferring psycholinguistic properties of words. In *Proceedings of the 2016 conference of the north American chapter of the Association for Computational Linguistics: Human language technologies* (pp. 435–440).

Parker, K. C. (1988). Speaking turns in small group interaction: A context-sensitive event sequence model. *Journal of Personality and Social Psychology, 54*(6), Article 965.

Paterson, K. (2023). Which format do the Most popular podcasts use?. available at https://www.thepodcasthost.com/business-of-podcasting/popular-podcast-format/. last accessed June 5, 2024.

Pew Research Center. (2025). Podcasts and News Fact Sheet. available at https://www.pewresearch.org/journalism/fact-sheet/podcasts-and-news-fact-sheet/?cb_viewport=tablet/. last accessed February 3, 2026.

Prinsloo, E., Scopelliti, I., Loewenstein, G., & Vosgerau, J. (2026). Responses to outcome disclosure: People asymmetrically disclose or hide their outcomes to protect others' emotions. *Organizational Behavior and Human Decision Processes, 193*, 1–24.

Rajar. (2020). Measurement of Internet Delivered Audio Services. https://www.rajar.co.uk/docs/news/MIDAS_Spring_2020.pdf. last accessed December 12, 2023.

Ratneshwar, S., & Chaiken, S. (1991). Comprehension's role in persuasion: The case of its moderating effect on the persuasive impact of source cues. *Journal of Consumer Research, 18*(1), 52–62.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*(4), 364–382.

Reid, S. A., & Ng, S. H. (2000). Conversation as a resource for influence: Evidence for prototypical arguments and social identification processes. *European Journal of Social Psychology, 30*(1), 83–100.

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review, 1*(1), 60–85.

Rime, J., Pike, C., & Collins, T. (2022). What is a podcast? Considering innovations in podcasting through the six-tensions framework. *Convergence, 28*(5), 1260–1282.

Robertson, S. (2004). Understanding inverse document frequency: On theoretical arguments for IDF. *Journal of Documentation, 60*(5), 503–520.

Robiady, N. D., Windasari, N. A., & Nita, A. (2021). Customer engagement in online social crowdfunding: The influence of storytelling technique on donation performance. *International Journal of Research in Marketing, 38*(2), 492–500.

Rocklage, M. D., Rucker, D. D., & Nordgren, L. F. (2018). The evaluative lexicon 2.0: The measurement of emotionality, extremity, and valence in language. *Behavior Research Methods, 50*(4), 1327–1344.

Rühlemann, C., & Gries, S. (2015). Turn order and turn distribution in multi-party storytelling. *Journal of Pragmatics, 87*, 171–191.

Sande, J. B., & Ghosh, M. (2018). Endogeneity in survey research. *International Journal of Research in Marketing, 35*(2), 185–204.

Schumann, K., Zaki, J., & Dweck, C. S. (2014). Addressing the empathy deficit: Beliefs about the malleability of empathy predict effortful responses when empathy is challenging. *Journal of Personality and Social Psychology, 107*(3), 475.

Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty, 26*(1), 5–16.

Statista. (2025). Podcast advertising in the United States - statistics & facts. available at https://www.statista.com/topics/7990/podcast-advertising/?srsltid=AfmBOops1SnxIakUyaovkpFF1qcx-xbNEZnvGBEwB6I0hC7alWYeLD1Z#topicOverview. last accessed January 21, 2026.

Stocking, G., Matsa, K. E., Naseer, S., Shearer, E., Jurkowitz, M., & Ghosh, S. (2023). Podcast format. Pew research Center, available at https://www.pewresearch.org/journalism/2023/06/15/podcast-format/, last accessed June 4, 2024.

Swiatek, L. (2018). The podcast as an intimate bridging medium. In *Podcasting: New Aural Cultures and Digital Media* (pp. 173–187).

Talbot, D. (2024). The New York times podcast statistics. available at https://wordsrated.com/the-new-york-times-podcasts/. last accessed June 25, 2024.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin, 110*(1), 67–85.

Valentini, C., & Kruckeberg, D. (2012). New media versus social media: A conceptualization of their meanings, uses, and implications for public relations. In S. Duhe (Ed.), *New media and public relations* (pp. 3–12). New York: Peter Lang.

Villarroel Ordenes, F., Grewal, D., Ludwig, S., Ruyter, K. D., Mahr, D., & Wetzels, M. (2019). Cutting through content clutter: How speech and image acts drive consumer sharing of social media brand messages. *Journal of Consumer Research, 45*(5), 988–1012.

Villarroel Ordenes, F., Ludwig, S., De Ruyter, K., Grewal, D., & Wetzels, M. (2017). Unveiling what is written in the stars: Analyzing explicit, implicit, and discourse patterns of sentiment in social media. *Journal of Consumer Research, 43*(6), 875–894.

Villarroel Ordenes, F., Packard, G., Hartmann, J., & Proserpio, D. (2025). Using traditional text analysis and large language models in service failure and recovery. *Journal of Service Research, 28*(3), 382–388.

Wang, X., Bendle, N., & Pan, Y. (2024). Beyond text: Marketing strategy in a world turned upside down. *Journal of the Academy of Marketing Science, 52*(4), 939–954.

Weiger, W. H., Moe, W. W., Wetzel, H. A., & Hammerschmidt, M. (2025). Do social media fans walk their talk? The impact of advocacy and criticism on own purchasing. *International Journal of Research in Marketing, 43*(1), 190–211.

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology, 50*(2), 314–321.