

The Mechanics of Auditory Stream Segregation in Polytemporal Media: A Systematic Literature Review of Cognitive Load and Attentional Selection

Dielan Wu ^{1*}

¹ Dr., Institute of Musicology, Central Conservatory of Music, Beijing, China

* Corresponding Author: wp24091@ccom.edu.cn

Citation: Wu, D. (2026). The mechanics of auditory stream segregation in polytemporal media: A systematic literature review of cognitive load and attentional selection. *Review of Communication Research*, 14, (Special)–93-109.

<https://doi.org/10.52152/RCR.V14.S7>

ARTICLE INFO

Received: 10 May 2026

Accepted: 15 Jul 2026

ABSTRACT

The polytemporal media requires the listener to follow more than one temporal order. This systematic literature review analyzes the role of auditory stream segregation in the facilitation, failure and overload of attention in the context of concurrent time structures in media soundtracks that contain speech, music, rhythm, ambience and interface cues. The goal is to bridge recent research in the fields of auditory scene analysis, cognitive load and selective auditory attention with communication and media research, where layered sound is a key element of communication but is seldom examined as a perceptual mechanism. Moreover, a systematic literature review method was used for the review and themes were analyzed through thematic analysis. Multidisciplinary databases, publisher platforms and open access indexes were searched, limiting the search to human listening studies and review articles that were relevant to stream segregation, attentional selection, listening effort and complex auditory scenes. The inclusion articles were 16 open access or openly available journal articles from the period 2019-2025. Four related mechanisms were identified in the review. The principal conclusion is that polytemporal media are aesthetically dense. They are auditory organization tests. A good design will direct the formation of streams, while a bad design will attract attention to fix the scene.

Keywords: Auditory Stream Segregation, Polytemporal Media, Cognitive Load, Attentional Selection, Communication Psychology.

INTRODUCTION

One pure stream of sound is seldom heard by a modern listener. The speech, music, room tone, alerts and caption rhythm are all carried in a news video. Speech, music, room tone, alerts and caption rhythm are all carried in a news video (Engstrom & Stephenson, 2024). A game scene encompasses dialogue, footsteps, weapons, wind, user interface sounds and adaptive music all in one listening field. A film score can have two tempos playing simultaneously, and voices can be moving around in space. A podcast may use music under narration. A social video is a video that is cut short, features meme sounds, and includes speech fragments that come before the last one has been processed. These are not details of the production. These are conditions for communication.

The concept of "polytemporal media" is used to describe a type of media architecture in which several time streams run concurrently, compete with one another, or are structurally connected. Polytemporalism is not just a densification of the senses, but a multi-level time ecosystem, in which different rhythms of operation coexist (Kurzyński, 2025). Speech has its own linguistic rhythm, background music its own rhythm, visual clips their own rhythm, and environmental cues their own rhythm in the same environment. This multi-level approach to the analysis of media allows to move the focus of analysis from the overall sensory intensity to the specific cognitive challenges, namely, the ability to follow several independent and parallel information streams simultaneously in

real time.

Communication research about complex acoustic environment analysis still has a key limitation. The evaluation of attention, audience reception, information design, media effects, persuasiveness and the changing nature of the audio technologies remains an integral part of this field of research, but the analytical approach of auditory flow separation has been rather neglected. The perceptual mechanisms to separate multiple streams of information in complex sound mixtures have been studied extensively by cognitive psychologists and auditory neurologists. Media scholars, however, tend to discuss layered sounds simply as a style, a part of the atmosphere, a layer of the “event” under way, without having to explain the perceptual processes.

Auditory stream segregation is the ability to hear a violin line in the orchestra, a voice in a party, or a notification amidst a music bed. It also provides an understanding of why certain sounds in the media will sometimes become noise. The listener has to do two things: group together acoustic events that are related and separate unrelated events. The auditory system uses various factors to determine what constitutes a stream, such as frequency, timbre, onset timing, spatial position, rhythmic regularity, and cue interaction (Hermes, 2023). In addition, such system is put under pressure by polytemporal media. They tend to use the same cues which enable stream segregation. A soundtrack can be used to give a center a narrator, sides ambience and music a lower dynamic range (Blauenfeldt, 2025). It can be used to distinguish icons from speech using timbre. It can synchronize visual cuts with sonic accents. It can also be the other way around. It can superimpose speech on top of a song with lyrics, add alert tones to the same pitch as speech, or keep the beats unrelated to the speech, distracting from the intended message. It is not just a matter of aesthetics that is in disorder. It is cognitive overload.

Communication researchers have a language to describe this pressure, cognitive load. Media design that forces the listener to search, switch, or repair unnecessarily consumes working memory, which means that the listener is not able to process the message (Brockhoff et al., 2023; Que et al., 2025). Theories of attentional selection include a second language. Attention can be directed by task goals, but can also be captured by salient events, familiar voices, sudden onsets and meaningful cues (Noyce et al., 2023; Rennig, Wegner-Clemens, & Beauchamp, 2020). The load and selection are not distinguishable in polytemporal media. The listener chooses, and the scene is more or less loaded depending on the success of the choice. The main issue that this review will focus on is shown in **Figure 1**. Polytemporal media are located between the acoustic organization, attentional selection and cognitive load. Communication happens when the listener is able to keep a target stream and there is sufficient context to allow meaning to be understood. When stream cues are too strong, the listener is forced to choose between comprehending the stream and attending to the environment, failure occurs.

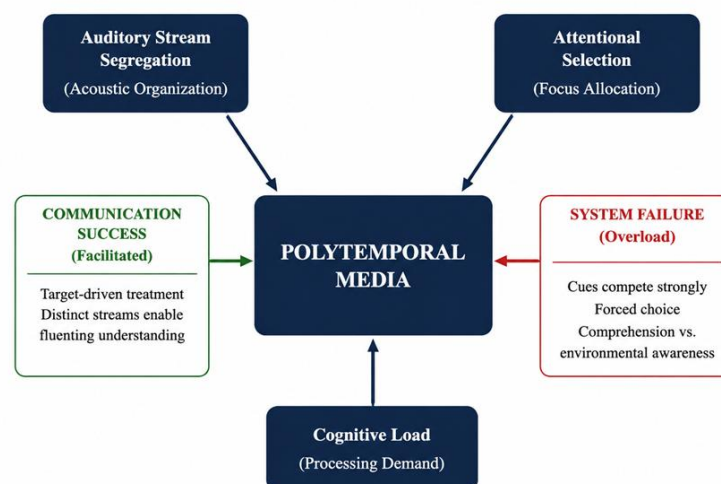


Figure 1. Conceptual Relation Among Polytemporal Media, Auditory Stream Segregation, Attentional Selection, and Cognitive Load

The review aims to have three objectives. The first is to combine recent open access research on auditory stream segregation and attention into a format that is accessible to communication scholars. The second is to relate stream segregation to cognitive load, particularly in environments with lots of sound. The third is to pinpoint the lack of evidence in the complexity of polytemporal media. The review is systematic in search and

screening, critical in interpretation. It is not included in a meta-analysis because the studies included are based on different stimuli, tasks, measures and theoretical frames

LITERATURE REVIEW

Auditory Stream Segregation as Perceptual Organization

Auditory stream segregation starts with a problem. The ear is not given a list of sources, but rather a mixture. The sound waves from voices, instruments, rooms, machines and devices all reach the eardrum at the same time. The auditory system needs to make inferences about the relationships between acoustic events. Recent research on auditory scene analysis (ASC) considers this to be an ongoing grouping task, in which listeners categorize competing sounds into perceptual objects and streams (Calculus, 2024; Gurariy et al., 2021). The concept is still basic, as it shifts the focus of listening research from individual tones to structured experience.

Repeated tone patterns (ABA sequences) are commonly used in streaming experiments. If the A and B tones are sufficiently different in pitch or quality and if they are presented at a sufficiently high rate, the listener will no longer hear one alternation. The sequence branches into two branches. The consequences of this simple demonstration are great. It demonstrates that the signal is not the only one that has temporal order. It is made by clustering. A listener might be misled about the rhythm, order or continuity due to events being allocated to different perceptual streams (Brown & Strand, 2019; Thomassen et al., 2022).

There are a number of acoustic cues that help to guide the process. The more frequently the two groups are separated, the greater the segregation. Timbre is used to indicate whether events are from the same source or not. The fusion is promoted by common onset and the separation is promoted by asynchronous onset. Space is one of the many cues that can separate sources. Recent pitch-timbre studies also reveal that cue relations can be compatible or incompatible, and stream segregation is not only based on one acoustic variable, but also on cue patterning (Oh et al., 2022; Roberts & Haywood, 2023; Jhang et al., 2025).

The idea is important for communication because messages are seldom stand-alone. Music, effects, ambience and other speech are competing for attention with speech in media. If these elements are well organized, listeners have richness without confusion. If cues are not well differentiated, the listener has to exert more effort to determine which cue to follow. The auditory scene is then a communication issue, rather than a sensory issue.

Cognitive Load and the Cost of Organizing Sound

The basic principle of cognitive load theory is applicable to mediated listening, and was developed in instructional psychology. There is a limit to the amount of information that working memory can handle at any given time. Design can aid learning and understanding by minimizing the need for unnecessary search and integration, or it can require too much mental effort from the audience to repair the poorly arranged materials (Brockhoff et al., 2023; Que et al., 2025). In AV communication, sound can be used to either decrease the load (attention cueing) or increase the load (dual task).

One thing is different about auditory load and visual clutter. Sounds begin and end. When a visual object is displayed, a verbal clause or alert tone is not displayed on the screen unless it is retained, replayed, or predicted (Mobin, 2025). Polytemporal media thus require high demands on the temporal working memory. The listener has to remember what has been said, predict what might follow and listen for the relevance of other streams. When lyrics are present in the same language as the narration, there can be competition between the two for language processing. When a rhythmic layer is in conflict with the prosody of speech, listeners might expend effort distinguishing between the two. A notification cue that looks like a musical accent might not be heard or misinterpreted as a musical accent.

Auditory-attention research provides an attentional account of recent research. The strength of task-irrelevant auditory events can become diminished or delayed in processing, or even completely missed, depending on the strength of the event and the listener's current task (Brockhoff et al., 2023; Noyce et al., 2023). Both types of pressure are frequently found in auditory media. The dense audio track can take up perceptual capacity, and the story task can take up executive control. The listener might then lose salient streams or get caught up in salient but irrelevant streams. This is supported directly in the literature in the form of multimedia and background music. Recent studies have found that background music can influence reading behaviour, and eye-movement, electrodermal and heart-rate measures can be used to predict various forms of cognitive load when reading with background music (Que & Hu, 2025; Que et al., 2025). Auditory stimuli that are not relevant to the task can also become undetectable when there is a working-memory load, particularly when attention is already focused elsewhere (Brockhoff et al., 2023). The results are consistent with the current music cognition research on

stream formation, in which the musical lines can split or merge based on pitch, timbre, working memory, listener expertise, and temporal order (Oh et al., 2022; Liu et al., 2025).

Attentional Selection in Complex Auditory Scenes

The cocktail party problem is used to explain selective auditory attention. One voice is heard by a listener and other voices are heard. Recent theories on auditory attention view attention as a process that is dynamic. It can improve an attended stream, inhibit competitors, and facilitate bottom-up salience and top-down goals (Noyce et al., 2023; Rennig, Wegner-Clemens, & Beauchamp, 2020).

There is no separation of attention and stream segregation. A listener might require segregation prior to selection of a stream, but attention can also help to stabilize a stream after it is selected. Speech and music provide evidence that attention enhances the attended object and can make ambiguous scenes more coherent (Gnanateja et al., 2022). Attention is a user resource and a design goal in the media. The sound designer's job is to direct the listener's focus by manipulating mix level, spatial position, spectral isolation, repetition, story value and timing. Audiences direct attention via goals, expectations, familiarity, language knowledge and previous media literacy.

This interaction is particularly apparent in polyphonic music and in scenes with a lot of sound in the audiovisual field. The lines of music can be heard separately or as a blended texture. The same principle is used in the audio for film and game. A leitmotif can transcend ambience as a result of its timbre, register and repetition. An interface cue can be separate due to its consistent envelope and spatial signature (Edina Fintor, Lukas Aspöck, Fels, & Schlittmeier, 2021). When this is the case, attentional selection is based on acoustic affordances embedded in the media object.

Polytemporal Media as a Communication Setting

Polytemporal media is not synonymous with polyphony. The term polyphony typically applies to musical lines that are played at the same time. Polytemporal media are any mediated scene that has more than one temporal stream that is relevant to meaning. A scene in a movie can have an ambient drone that is slow, and then have fast dialogue. A documentary can include narration, archival audio, music and subtitles. A live stream can include voice chat, gameplay effects, music, and platform notifications. The listener is asked to localize the environmental sounds in a virtual reality scene, while listening to verbal instructions. It is not genre that is the common denominator. It is a simultaneous demand in time.

This concept can be applied to the communication research to connect media studies and cognitive listening research. Reception is not interpretation after perception. It starts with perceptual organization. If a stream does not become a stream, then a listener cannot interpret it. Nor can a listener listen to all streams at the same time if the design calls for incompatible time commitments. Polytemporal media thus pose a fundamental question: When is richness overload?

The key concepts of this review are defined in **Table 1**. The table is presented early since the literature employs overlapping terms and a common language is required for synthesis to take place. Table 1 illustrates that auditory stream segregation, attentional selection and cognitive load refer to different aspects of a single listening event.

Table 1. Core Concepts used in the Review

Concept	Working Definition	Media Relevance
Auditory stream segregation	The perceptual grouping of acoustic events into one or more coherent streams or auditory objects.	Explains how speech, music, ambience, and alerts become separable or fused in a soundtrack.
Polytemporal media	Media environments in which two or more temporal streams coexist and require perceptual or interpretive handling.	Includes films, games, podcasts, social video, VR, live streams, and multimedia learning.
Cognitive load	The demand placed on limited working memory by intrinsic task complexity and unnecessary design burden.	Explains why layered sound can support meaning or exhaust attention.
Attentional selection	The prioritizing of one stream, object, location, feature, or message over competitors.	Explains how audiences follow a target voice, instrument, cue, or narrative layer.
Auditory object	A perceptual unit inferred from acoustic features that seem to belong to one source.	Helps connect laboratory streaming with media sound design.

Auditory Stream Segregation in Communication Research

Auditory stream segregation is a theory of how listeners make an ordered message field from competing signals (Gustafson, Nelson, & Silcox, 2023). This is a communication issue, not a lab issue. Listening to news, watching a film, going to a theatre, watching a public warning system, listening to a radio broadcast, playing a game, watching an educational video, listening to a podcast and using a mobile interface all rely on listeners correctly matching sounds to a source at the right time. If that assignment is not achieved, the message isn't just less pleasant. It is less available.

When communication is considered a scarce public good, the value of the evidence becomes more apparent. Media not only vie for screen time, but perceptual priority (Kankam, 2025). A voice-over should be given priority over music. The warning tone is to be given priority over background noise. A game cue should not be a distraction to the game. In the narration of a teacher in the multimedia material, the narration should be coherent as the images and sounds are presented. Typically these cases are examined in terms of design, persuasion, usability or reception. A mechanism is added under them, called stream segregation.

The idea also prevents two pitfalls. The first error is technological determinism, which assumes that the richer the media, the better the communication. The first error is technological determinism which assumes that the more channels the media has, the better the communication. The second is simplification, which is when designers remove sound from the design since complexity is seen as a risk. There is no evidence here to back either side. It demonstrates that having multiple streams can enhance communication when grouping cues, hierarchy, and listener goals are aligned. It also demonstrates that a small amount of music can be distracting when the streams are not well separated or when they are incorrectly grouped.

Therefore, polytemporal media should be studied as listening environments which are organized. They are not simply a means of conveying messages. They are perceptual structures which facilitate the selection of some messages and not others. In this review, therefore, auditory mechanics is considered to be a part of communication mechanics (Gustafson, Nelson, & Silcox, 2023). It is a subtle change in the language, but a significant one in terms of impact. It brings the analysis of the media to the forefront of the audience's experience.

Moreover, the goal of this study is to find out mechanisms and to translate them into a communication research frame. Moreover, the review was based on three questions. So, what are the most important mechanisms of auditory stream segregation for polytemporal media? Second, what are the interactions between attentional selection and stream formation in complex auditory scenes? Third, what is the cognitive load involved in listening to multiple temporal streams when the media sound needs to be followed, suppressed or integrated?

METHODOLOGY

Review Design

A systematic literature review was used as a method in this study. It did not include review and interviews, experiments, bibliometric mapping, or meta-analysis. The field is dispersed in the areas of auditory perception, cognitive neuroscience, music cognition, media psychology, human communication, and multimedia learning, so a systematic review was selected. The intent was not to calculate an overall effect size. Too much variation in studies with respect to stimuli, listeners, measures, and outcomes for that approach. Furthermore, the thematic analysis was applied to systematically break down and combine different findings from various fields, rather than a mere chronological review. Thematic analysis was used to identify underlying structural patterns, as the literature included in the book ranged from highly controlled auditory science to music cognition and applied communication studies. This conceptual approach, which judged each study by its content, rather than its experimental design (Squires, 2023), was effective in synthesizing a disjointed landscape into four coherent, mechanism-based themes. Finally, it made what was once a collection of laboratory findings into a unified approach to sound design and media theory in practice.

Search Strategy and Inclusion Criteria

The literature search included peer-reviewed journal articles on the topics of auditory stream segregation, auditory scene analysis, selective auditory attention, listening effort, cocktail party perception, polyphonic music attention, and cognitive load in AV or multimodal communication. Search strings included auditory stream segregation, auditory streaming, auditory scene analysis, selective auditory attention, cocktail party, cognitive load, listening effort, polyphonic music, naturalistic auditory streams, and multimedia sound. A search was carried out in Google Scholar, Web of Science, Scopus, PubMed, Crossref, DOAJ (where applicable), publisher search pages and open access repositories. From the central review articles and highly cited studies, reference chaining was then used.

Articles were selected if they fulfilled five criteria. They were required to be from an academic journal article. They were required to be available as open access or openly available full text in a publisher, PubMed Central, institutional repository or open journal. They were required to solve the problem of human auditory stream segregation, auditory attention, listening effort, or related mechanisms in complex auditory scenes. They had to be relevant to media environments where multiple temporal streams compete or cooperate (Ragert, Fairhurst & Keller, 2014). They were required to be in English. Articles were discarded if they were purely computational, clinical-only with no relevance to attention or media listening, were only about hearing device engineering without perceptual findings, or were not available as a full text.

A total of 16 articles were obtained after the screening. All sources retained were published between 2019 and 2025, and were openly accessible or open access via a publisher site, repository, or open journal platform. The screening logic for the review is presented in **Figure 2**. The search and screening decisions are summarized in **Table 2**.

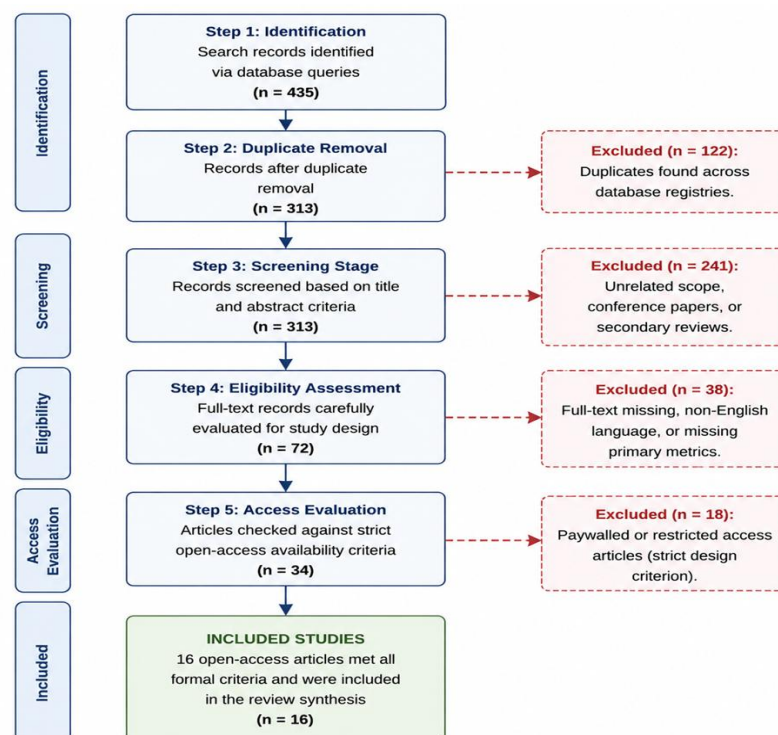


Figure 2. Systematic Literature Review Screening Pathway From Search Records To 16 Included Open-Access Articles

Table 2. Search Strategy and Screening Criteria

Stage	Decision Rule	Outcome for this Review
Identification	Search multidisciplinary databases and publisher platforms using stream segregation, auditory attention, listening effort, and media sound terms.	Records were gathered across psychology, neuroscience, hearing science, music cognition, and communication-adjacent literatures.
Deduplication	Remove repeated records and non-journal items.	Books, theses, preprints without journal publication, and conference papers were not counted in the corpus.
Title and abstract screening	Keep studies that linked auditory segregation, selective attention, load, or complex auditory scenes.	Studies limited to single isolated tones without interpretive relevance were deprioritized unless theoretically central.
Full-text eligibility	Keep open-access or openly available academic journal articles.	Exactly 16 articles were retained for systematic synthesis.
Synthesis	Code each article for stimulus type, listener task, mechanism, attention role, load relevance, and media implication.	A qualitative thematic analysis was used instead of a meta-analysis.

Review and Coding Procedure

The six features that were coded for each article were domain, stimulus type, segregation cue, attention mechanism, load implication, and communication relevance. The article was categorized as primarily belonging to psychoacoustics, neuroscience, music cognition, speech perception, media effects, or multimedia learning. The stimulus type distinguished between tones, speech, music, natural sound scenes and audiovisual materials. The primary basis for stream separation (frequency, temporal coherence, timbre, onset, rhythm, or spatial location) was segregated. The attention mechanism was implemented as either selection, neural gain, task instruction, object formation, or executive control. Load implication was recorded as direct load, implied load due to performance limits, or design principles related to load. Communication relevance was coded as high when the findings were clearly relevant to media sound, speech reception, audience attention, immersive media or message design.

The 16 articles included are listed in [Table 3](#). The table also indicates the rationale for the cross-disciplinary nature of the review as opposed to a strictly auditory one. Polytemporal media are located at a border. They take from speech, music, environment and interface sound. There is no one research tradition that addresses the entire problem.

Table 3. Sixteen Included Open-Access or Openly Available Journal Articles

Included Article	Journal	Main Focus	Mechanism Coded	Use in Synthesis
Calcus (2024)	Frontiers in Human Neuroscience	Auditory scene analysis review	Segregation and selective attention	Updates core ASA concepts for recent stream-formation evidence.
Gurariy, Randall, & Greenberg (2021)	Psychological Research	Auditory object grouping	Low-level acoustic features	Supports cue-based grouping strength in sound-object formation.
Ferrario and Rankin (2021)	Journal of Mathematical Neuroscience	Auditory streaming model	Excitation, inhibition, streaming dynamics	Explains dynamic transitions between integrated and segregated percepts.
Brown and Strand (2019)	Journal of Cognition	Effect of visual talker cues on speech recognition and listening effort.	Audiovisual integration; visual speech cues; increased cognitive listening effort	Shows visual cues improve recognition but increase cognitive load.
Rennig, Wegner-Clemens, & Beauchamp, (2020)	Psychonomic bulletin & review	Audio-visual speech information	Infrared eye tracker and audiovisual through Matlab, and Psychophysics Toolbox	Multisensory gain during speech.
Oh, Zuwala, Salvagno, & Tilbrook (2022)	Frontiers in Neuroscience	Pitch and timbre grouping	Pitch, timbre	Supports spectral and timbral differentiation as design cues.
Thomassen, Hartung, Einhäuser, & Bendixen (2022)	Journal of the Acoustical Society of America	Sequential auditory scene analysis	Pattern order, temporal sequence	Clarifies how tone patterns can alter segregation decisions.
Noyce, Kwasa, and Shinn-Cunningham (2023)	WIREs Cognitive Science	Auditory attention framework	Top-down and bottom-up attention	Defines attention as a dynamic auditory process for complex scenes.
Brockhoff et al. (2023)	Scientific Reports	Working-memory load and auditory stimuli	Cross-modal load, awareness	Links cognitive load with detection of task-unrelated sounds.
Que and Hu (2025)	Reading Research Quarterly	Background music, reading comprehension, and cognitive engagement	Auditory stimulation; cognitive attention regulation; physiological response during learning	Background music can influence comprehension through changes in attention and cognitive processing.
Roberts & Haywood (2023)	Journal of the Acoustical Society of America	Timbre-change effects	Timbre changes	Shows timbre shifts can asymmetrically affect stream segregation.

Included Article	Journal	Main Focus	Mechanism Coded	Use in Synthesis
Vosburg (2024)	Review of Communication Research	Media effects and visual communication review	Information processing, recall	Adds communication-theory context for multimodal media effects.
Jhang, Ueda, Takeichi, Remijn, & Hasuo (2025)	PLOS ONE	Pitch-timbre rivalry	Pitch and timbre trade-off	Updates evidence on cue competition in sequential stream segregation.
Liu et al. (2025)	Frontiers in Psychology	Musicianship and stream segregation	Auditory working memory	Connects listener differences and memory mechanisms to segregation.
Que, Zheng, Hsiao, & Hu (2025)	Scientific Reports	Measuring cognitive load with background music	Eye movement, EDA, HR	Gives current non-intrusive load measures for layered audio contexts.
Yang, Sannusi, & Ahmad Rizal (2025)	Review of Communication Research	Immersive journalism meta-analysis	Presence, interactivity, information processing	Adds communication evidence on immersive media complexity and adoption.

Synthesis Approach

The synthesis was not chronological, but rather based on the mechanism. This was required because the literature does not advance in a linear fashion. Some recent reviews give a general conceptual framework, and newer empirical papers give very specific measures but only under very specific stimulus conditions. A few articles that are relevant to communication do not employ the terms of stream segregation. The review thus grouped the evidence into four themes: acoustic grouping, attentional stabilization, cognitive load and media translation.

The evidence was not evaluated on a statistical basis, but rather for conceptual contribution. This decision ensures that the review remains in line with the journal's interest in comprehensive critical reviews. It also prevents a false precision that would result from combining incompatible designs. The review is systematic with regard to the sources selected and transparent with regard to coding, but interpretive with regard to the results.

RESULTS

The thematic analysis revealed four interdependent mechanisms of listeners' processing of complex polytemporal media. First, the baseline sensory organization is Acoustic Cues (frequency, timbre, temporal coherence) which makes stream segregation easier or harder. Second, an Attentional Loop is identified, which shows that selection and stream stability mutually reinforce during active media consumption. Third, Cognitive Load is not just a function of the number of sounds in the soundtrack, but rather the mental effort needed to resolve the ambiguity and competing layers of sounds. Lastly, an Ecological Gap points to the lack of connection between the sterile laboratory tone tests and the actual communication environment, and requires a common design and theoretical language.

Theme 1: Acoustic Cues Create the First Listening Order

The first is that polytemporal media rely on cue management prior to relying on interpretation. Acoustic evidence is used by the auditory system to determine what goes together. Frequency separation is still one of the best indicators in streaming. If two repetitions of the same tone are presented at a high frequency difference and high repetition rate, listeners will hear two streams instead of an alternating pattern (Brown & Strand, 2019; Thomassen et al., 2022). Timbre is similar. Instruments with different timbral envelopes are easier to follow as separate lines than instruments with similar timbral envelopes.

The time of onset is also important. Sounds starting at the same time will tend to merge. Sounds that start at different times are more likely to segregate. This is why a music cue can mask a speech onset if both are presented at similar levels, and why a clean onset for speech can make speech intelligible even in the presence of background music. Spatial position is a good aid, but it is not a cure. A voice that is slightly separated from music can become more audible, but spatial separation can be overcome by loudness, spectral masking, reverberation, or task requirements.

Temporal coherence provides a helpful link to media. Recent studies indicate that listeners' retention of streams over time is influenced by pitch, timbre, frequency-band structure, and abrupt acoustic changes (Oh et al.,

2022; Roberts & Haywood, 2023; Jhang et al., 2025). This is the same for the design of audiovisuals. Temporal coherence should be shared by a group of sound effects that are associated to a single visual object. The temporal signature of an interface sound should be stable. Ambience should be used to complement the scene, but not to mimic the modulation of speech. Polytemporal media are manageable when the coherence of each layer is sufficient to be heard as a layer. This first order mechanism is shown in **Figure 3**. Attention is not determined by acoustic cues, but rather is easier or harder to pay. When the listener cannot tell the difference between the cues, he or she must solve a grouping problem before the message can be understood.

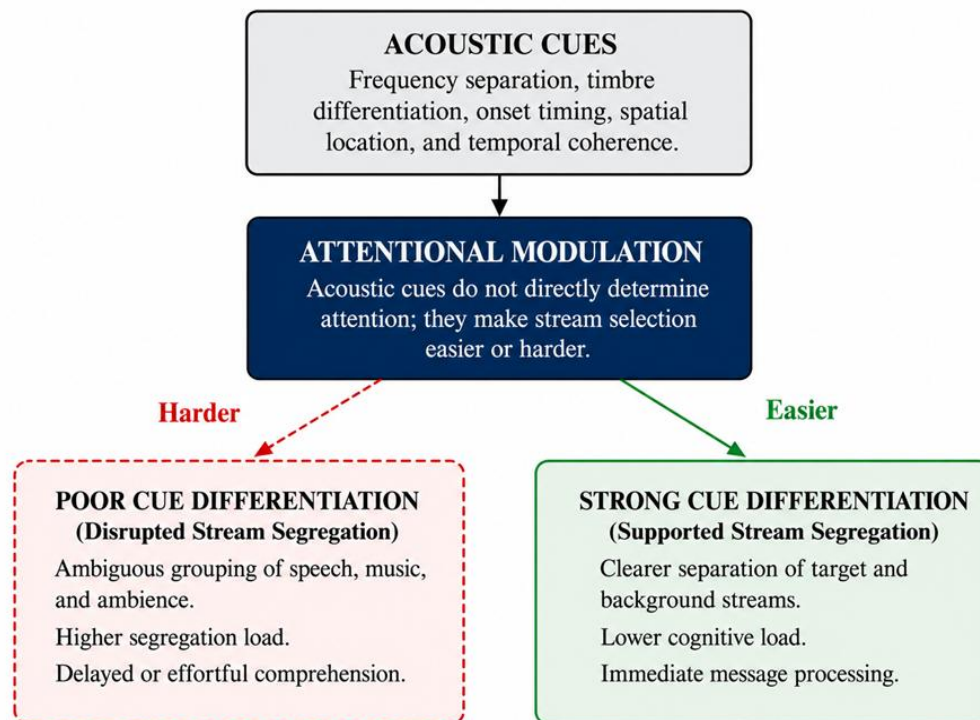


Figure 3. Mechanisms By Which Acoustic Cues Support Or Disrupt Auditory Stream Segregation In Polytemporal Media

Theme 2: The Mutual Dependency of Attention and Stream Stability

The second finding is that there is a circular relationship between segregation and attention. Perceptual organisation is required for listeners to select a target, but selection can also enhance the target stream once it has been selected. This is evident in the studies on selective speech and audiovisual speech. Selective attention to AV speech can be enhanced by spatial alignment of faces and voices, in addition to sound alone (Rennig, Wegner-Clemens, & Beauchamp, 2020; Noyce et al., 2023).

Recent auditory-attention studies demonstrate that the attended stream is measurable in behavior and neural responses in complex listening. The discovery is discussed in the field of hearing technology and cognitive neuroscience, but has a much wider scope of communication. Media listeners don't passively listen to a mix. They are in the process of stabilizing one part of it (Noyce et al., 2023; Brockhoff et al., 2023).

Music cognition follows the point. Liu et al. (2025) connect auditory working memory, musicianship and auditory stream segregation, demonstrating that listeners' resources and experience influence the organization of musical materials (Symonds et al., 2019). This discovery is important as it brings the literature closer to polytemporal media. Music does not provide streams that are isolated speech channels. It provides logical, but connected lines. When there is more than one line, attending to one line alters what the listener hears as structure.

The review corpus therefore does not support a simple model of sound design, which generates streams and then selects one. In real listening, the two processes are "linked". The design is inviting, attention confirms that, expectation expects it, and load either confirms it or destroys it. **Table 4** shows the coding system used to identify this loop in the articles included.

Table 4. Coding Framework Used in the Synthesis

Code	Question asked of Each Article	Pattern Observed
Stimulus ecology	Were stimuli tones, speech, music, natural scenes, or multimedia materials?	Tone studies clarified mechanisms; speech and music studies carried stronger media relevance.
Segregation cue	Which acoustic property organized streams?	Frequency and timbre dominated classic work; temporal coherence bridged laboratory and media contexts.
Attention role	Was attention a task instruction, neural gain process, object selection, or executive control factor?	Attention both selected and stabilized streams.
Load relevance	Was cognitive load measured directly or inferred from task pressure?	Direct load measures were rare; load was often implied by performance limits and competing demands.
Media transfer	Could findings apply to polytemporal media design?	Strongest transfer occurred in speech competition, polyphonic music, and multimedia learning.

Theme 3: Cognitive Load Appears When Listeners Must Repair the Scene

The third outcome is that the number of sounds is not the only measure of cognitive load. It shows up when the listener has to make unnecessary organization efforts. It might be easier to have a dense, but well-structured soundtrack than a sparse, but ambiguous one. This is significant for media studies as it is easy to confuse richness with complexity. When the listener cannot discern which streams are important, which can be disregarded, and how the streams are related over time, richness becomes a problem.

Audio layering research in the recent past demonstrates that layered audio can lead to extraneous effort when it is not helpful to the primary task. Que et al. (2025) consider background music while reading as a situation where various types of cognitive load can be predicted based on eye movements, electrodermal activity and heart-rate. If a sound layer is not appropriate for the task, it could be using resources that could be used for comprehension. This issue is exacerbated when music and speech have both semantic content. Lyrics that are narrated are not background. They can be a competitor for the verbal stream.

The second load pathway is added from the auditory attention literature. Competition requires listeners to ignore distractors and to keep track of the target. This is not free. Competing speech is hard because of recent auditory-attention frameworks, which highlight the shared linguistic structure, temporal modulations, and social relevance of voices (Noyce et al., 2023; Rennig, Wegner-Clemens, & Beauchamp, 2020). The same problem occurs in polytemporal media if multiple layers are salient in terms of their human-like timing or story.

Load also increases when there is a conflict between temporal scales. The listener can follow the speech at syllabic or phrasal speed, while the music is metrically in a different cycle, and the visual editing has its own speed. Comprehension can be enhanced if the streams are supportive. When they compete, the listener will have to switch, inhibit or reconstruct the object structure. The review thus does not consider cognitive load to be a straightforward measure of the density of the soundtrack, but rather an emergent property of the stream organization and task demand.

Theme 4: Polytemporal Media Expose the Limits of Laboratory Streaming

The fourth one is a gap. The strongest literature in the area of auditory streaming tends to be the controlled tone sequences, and the strongest problems in the area of communication are in the ecological media environments. The mechanisms that are at the heart of tone studies are revealed. They demonstrate frequency separation, repetition, timing, timbre and cue rivalry (Ferrario & Rankin, 2021; Thomassen et al., 2022; Jhang et al., 2025). However, they lack semantic attention, narrative expectations, visual synchrony, conventions of the platform, and user goals.

Speech studies are more ecological, but still are very much about the cocktail party case. That case is at the heart of it, but polytemporal media are more than the competition of talkers. These are: music in speech, asynchronous media cues, environmental sound, notification layers, game audio, voice-over, visual rhythm and user interaction. In a game, a listener can be required to listen to dialogue, localize a threat, listen to team chat, etc., and interpret the intensity of the music at the same time. The listener in a documentary might have to combine archival speech and narration and image evidence. These are problems involving the segregation of streams of communicative stakes.

There is also a disciplinary split, as revealed by the review. Auditory science is precise in its description of

mechanisms, but it does not often relate to communication theory. Layered sound, immersion and narrative are topics of discussion in media studies, but not often explained with the concepts of stream segregation or cognitive load. There is a need for a medium level vocabulary in the field. It can be given by polytemporal media. It enables scholars to consider how temporal layers are perceived prior to being interpreted culturally. The results of the study are summarized as design and theory implications in **Table 5**. It demonstrates that there is a risk in media environments for each mechanism and a research opportunity.

Table 5. Mechanisms, Media Risks, and Research Implications

Mechanism	Media Risk	Research Implication
Frequency and timbre segregation	Speech, music, and alerts overlap in the same spectral region.	Study mix design as perceptual organization, not only aesthetics.
Temporal coherence	Separate media elements move together by accident and are falsely grouped.	Analyze synchronization among sound, image, and task.
Attentional gain	Salient distractors pull attention from target streams.	Measure attention to media layers, not only recall of final messages.
Cognitive load	Listeners spend resources repairing ambiguous scenes.	Develop load measures for layered sound and temporal competition.
Predictive listening	Media violates expected sound object continuity without communicative purpose.	Connect auditory prediction with narrative expectation and interface feedback.

DISCUSSION

A Communication Model of Stream-aware Listening

The findings suggest a model of stream-aware listening: a communication model. This model is one in which meaning does not occur after perception. Meaning starts as listener's streams. If the target voice, cue or musical line does not turn into a stable auditory object, then the interpretation is already affected later on (Symonds et al., 2019). The process of communication research typically starts with the exposure, attention, comprehension and effect of the message. The present review proposes a previous step: perceptual object formation (Gurariy et al., 2021). This model is shown in **Figure 4**. Candidates are formed in the streams by acoustic cues (Jhang et al., 2025; Oh et al., 2022). Attention is focused on and maintains one or more of them. The more streams the listener has to keep in mind or fix up, the higher the cognitive load. The selected stream is sufficiently stable to allow memory, inference and emotional response. The model is simple, but it can be used to connect the auditory science and media research.

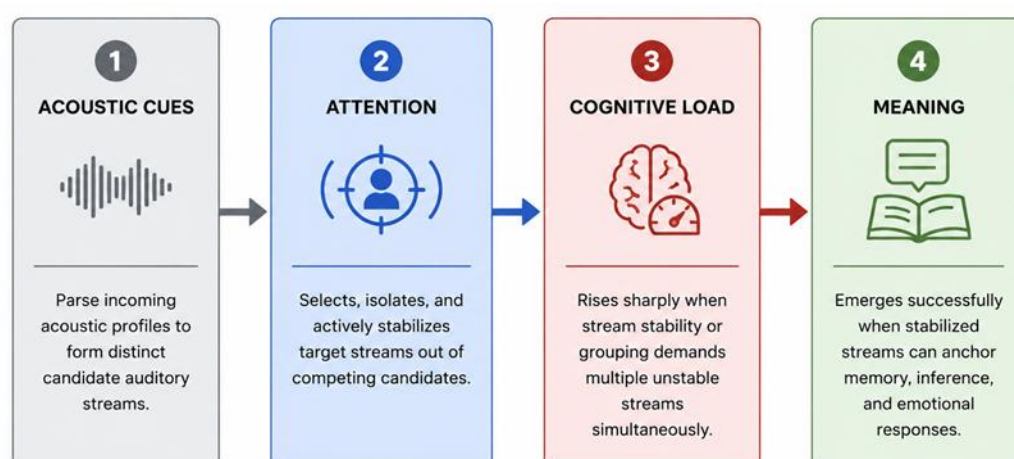


Figure 4. Stream-Aware Listening Model for Polytemporal Media Communication

This model also helps to explain the potential of polytemporal media. Not all multiple streams are necessarily bad. They enable irony, suspense, immersion, spatial realism, musical counterpoint and emotional layering. Multiplicity is not a problem. The issue is that of multiplicity that is not managed. A good soundtrack can distinguish streams sufficiently to allow attention to be paid to them, and sufficiently to allow them to be

interpreted. If the music is not good, the listener has to decide whether to listen or understand (Roberts & Haywood, 2023; Thomassen et al., 2022).

The model will influence the way that scholars will assess media sound. Loudness balance is important but it is not the only thing. Analysts should consider whether the target stream has a distinct onset, whether it is in a secure spatial region, whether its rhythm is at odds with competing layers, whether graphical information helps or hinders stream formation, and whether salient distortions occur at moments critical for understanding.

Cognitive Load as Temporal Management

The best understanding of cognitive load in polytemporal media is temporal management. The listener has to control what is going on, what has just occurred, what is to come and which stream is to be given priority. This is more complicated than attaching sound to a message. It is a type of time coordination.

There are three types of loads suggested by the review. When streams are difficult to separate, segregation load will be present. Selection load is when multiple streams are distinct and are competing for relevance. Integration load is when streams need to be integrated over time, such as when narration explains the archival sound or music alters the meaning of a visual scene. These forms may be merged together. A listener might have a difficult time distinguishing a voice from music, then have to determine if the music is important, and then merge the voice and music with an image sequence.

This is helpful because it helps to differentiate between the various design options that address various loads. Spectral spacing, clean onsets, spatial separation and dynamic range management can all help to reduce segregation load. Stable hierarchy and cueing can help to reduce selection load. Temporal alignment, repetition and meaningful redundancy are all ways of supporting integration load. Making art simple does not mean reducing the load. It involves the elimination of unintentional barriers to communication.

Attentional Selection is Guided, not Merely Captured

The review also questions the notion that listeners are directing the focus of their attention to wherever they want. Design leads the attention in polytemporal media. Selection is biased by salient onsets, familiar voices, high contrast timbres, spatial motion, lyric content, and narrative relevance (Ferrario & Rankin, 2021). Goals are set by the audience, and the soundscape offers affordances. That is why sound effects in the media can capture attention without being told.

The risk is of being distracted for no good reason. If there is a sudden sound, it can distract the listener from the dialogue. A high frequency cue can break up the continuity of emotion, but can also cut through a mix (Noyce et al., 2023). When the designer was looking for atmosphere, a familiar song under speech can evoke autobiographical or lyrical processing. There are a lot of traps in polytemporal media, as streams are both acoustic and meaningful.

The implication for communication research is methodological. Media reception studies should not just be based on the extent to which participants remember content. They should analyze which streams they followed and which they ignored and where their focus moved. The use of eye tracking is now ubiquitous in the study of visual media. There are still many methods that are sensitive to the attention of sound that are not used in communication studies, such as behavioral probes, response timing, pupillometry, EEG, and self-report mapping.

Toward a Research Agenda for Polytemporal Media

There are a number of gaps identified in the review. First, most studies of auditory stream segregation continue to use simplified stimuli. While this is required for mechanism, it is not enough for media theory (Calculus, 2024). More research needs to be conducted using naturalistic media sound, with experimental control. Second, in the studies of auditory attention, cognitive load is not necessarily measured directly. Media research can be helpful by creating tasks that will involve comprehension, recall, emotional response, and attention switching (Brockhoff et al., 2023). Third, music and speech are studied separately, whereas they are always combined in the media. Fourth, there are still insufficient differences in listeners. The organization of polytemporal sound can be affected by expertise, age, language familiarity, hearing status, media habits and musical training. A suggested framework for future communication research is given in [Figure 5](#). It connects media-layer analysis, listener-task analysis and perceptual mechanism analysis. For instance, a research on a game scene would record the sounds for each layer of the game, determine the listener's task, quantify stream tracking and mental effort, and correlate the results with comprehension or engagement. A podcast study would focus on the ways music beds, pauses, voice timbre and transitions in the narrative draw attention to different moments in time.

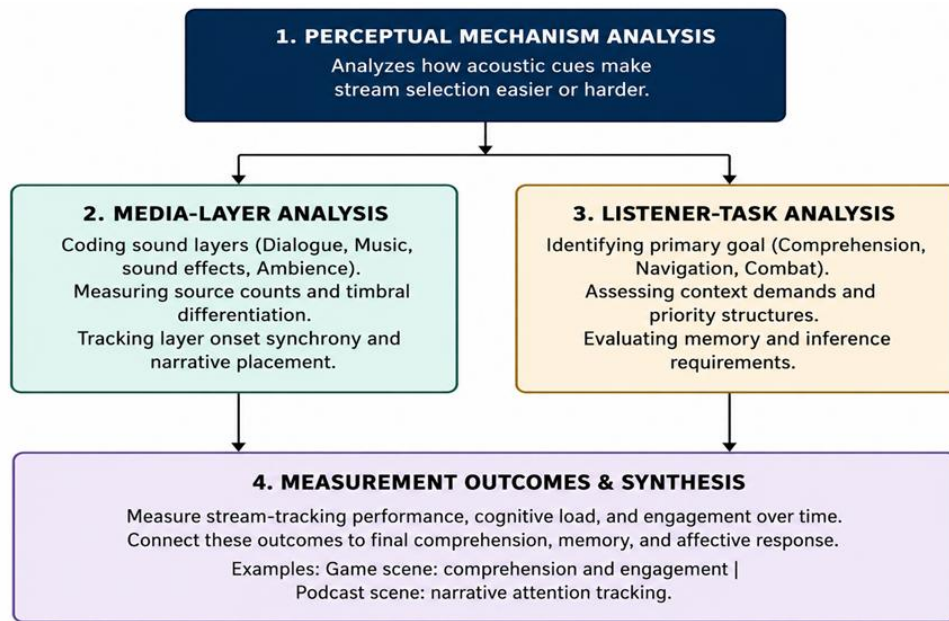


Figure 5. Proposed Research Framework for Studying Polytemporal Media, Stream Segregation, and Cognitive Load

This approach is in line with the current research trends, which combine communication psychology, media technology, neurocommunication, speech communication and audience research. The critical syntheses that have been carried out in the field in recent years illustrate the tremendous potential of linking media effects, information processing, and immersive media complexity in various communication contexts (Vosburg, 2024; Yang et al., 2025). Moreover, this framework extends the scope of the current literature that is confined to a laboratory or clinical setting. The auditory system is not only a passive sensor, but also an active communication interface in today's mediated environment.

The best that can be gleaned from the literature reviewed is not a single rule, such as, separate frequencies or lower background music. It is a relational principle. A media stream is communicative when it is perceptually different from others and is connectable to others. For designers and researchers, then, sound layers are to be considered as competing temporal claims for attention.

Design Implications for Polytemporal Media

The evidence reviewed provides a number of practical rules for communication design, however these should not be interpreted as a formula. A dense, but stable, role for each layer makes a soundtrack clear. There should be protected times at the beginning of significant clauses, particularly where new information is introduced, in the target stream. Music should not be in the same semantic and spectral space as narration, unless there is an intention to interfere. Alert sounds should not be the same timbre, envelope or spatial location as musical accents (Edina Fintor et al., 2021). Ambient sound should provide place and mood, but not mimic modulation pattern of speech. These are not just preferences for production. These are methods to lower segregation loads.

There is also a hierarchy of attention for polytemporal media. The hierarchy can shift from dialogue to music to environmental sound within a scene in film. It can be voice chat, threat cues, navigational feedback, etc. in games. In podcasts, voice is typically the most important element, but music can be used to indicate transition, effect, irony, or memory. The listener is able to follow the hierarchy when it is presented to them. When the hierarchy changes suddenly without warning, it is expensive since the listener has to give up one temporal object and build another. Such changes are not without significance, as the review corpus demonstrates. While attention can help to stabilize a stream, there is a cost when a previous stream is still active and meaningful.

The second implication relates to the accessibility. Listeners vary in terms of their hearing, language proficiency, musical background, age, fatigue and media consumption. The combination that works for one person might not work for another (Brown & Strand, 2019). Design that is stream-aware thus promotes inclusivity. It enables captions, visual rhythm, volume, spatialization and repetition to all play a part. It also provides an understanding of why accessibility should not be considered a post-production problem of captioning (Rennig et al., 2020). The soundscape can be more or less accessible depending on the distribution of temporal information across the streams.

Last but not least, listening literacy should be incorporated into media literacy. People are taught to read pictures and stories, but not to listen to structures of sound. A student watching a documentary can ask how the narration, archival sound, music, silence and image rhythm lead the viewer's eye. A researcher in game studies can ask about the player's sense of priorities when time is scarce. A health communication scientist may wonder if background music is a help or a hindrance to verbal instructions. The questions bring auditory stream segregation into the realm of ordinary communication analysis.

Methodological Implications for Future Reviews and Experiments

The review also highlights lessons to be learned from the methodology. First, future studies should differentiate between studies that measure stream segregation directly and studies that infer stream segregation based on speech comprehension, neural tracking or listening effort. These are similar, but not the same results. It is possible to be able to comprehend speech without having high load. A different listener can segment streams, but choose the incorrect stream. A third can follow a target stream, but without background emotional or contextual information. Communication research needs measures that are sensitive to these differences.

Secondly, experiments in the future should be conducted in the laboratory with media materials, but without losing control. Stimulus sets may be generated from short film clips, game scenes, podcasts, instructional videos, or social media types. One of the following features may be manipulated: background music with or without lyrics; aligned versus misaligned sound onsets; spatially separated versus centered narration; stable versus shifting tempo; or predictable versus unpredictable alert timing. These designs would allow researchers to explore the principles of streams and maintain the ecological richness that tone sequences do not provide.

Thirdly, reporting needs to be better. While many studies are sufficiently detailed to be replicated in auditory science, they are not sufficiently detailed to be understood by communication scholars to understand media relevance. Reports should contain data on the number of streams, the temporal rates, the overlap of the streams, the semantic content, the spatial placement, and the relevance for the task. The description of an audio-layer might be as useful as a description of a visual stimulus in media research. If it is not, it is difficult to transfer studies of auditory attention to other fields.

Fourthly, load measures should be layered. Self-report can be used to measure effort, but not rapid attentional shifts. Target tracking can be captured with behavioral probes, but can be disruptive to natural listening. The process is revealed in various ways, such as through pupillometry, EEG, response times, comprehension tests, and post-exposure mapping (Que & Hu, 2025). Such evidence can be synthesized in a systematic review, but it must be collected in experimental work in a manner that is sensitive to the temporal nature of sound. Polytemporal media requires methods that can be used to guide the listener in time, not only before and after the listener is exposed to the media.

CONCLUSION

The systematic literature review explored the issue of auditory stream segregation in polytemporal media in relation to the related issues of cognitive load and attentional selection. A total of 16 open access or openly available journal articles were systematically reviewed and foundational sources were used to define the older theories that are still used in the field. Polytemporal media are not only complicated soundtracks, as the review demonstrates. They are communication contexts which challenge the listener's temporal organization.

There are four conclusions to note. First, the acoustic cues that stream segregation starts with are: frequency, timbre, onset timing, spatial location, and temporal coherence. Second, attention is not a late stage option, after segregation. It stabilises streams and can modulate the representation of the auditory scene. Third, listeners' cognitive load increases when they have to overcome the lack of stream organization, attractive distractors, or multiple temporal layers without clear hierarchy. Fourth, the existing research literature is not yet up to the challenge of the complexity of real media sound. Tone sequences and cocktail party speech still play a key role, but are not fully representative of games, films, podcasts, social video, virtual environments, and multimedia learning.

The article provides a helpful connection between the fields of communication and research. It links sound science with media design and audience perception, and also provides a direct guideline for sound designers and media scholars; to take the lead in constructing the "sound flow" (stream), rather than passively waiting for it to form. Unorganized sound cannot be processed by the auditory system. Excellent "polytemporal" media do not eliminate complexity; instead, they make this complexity audible.

Limitations and Future Research

The review has some limitations. It was limited to 16 open access or openly available journal articles, making

the review transparent but also selective. The topic spans across a number of disciplines and some important closed access studies were not included in the systematic corpus. The synthesis is qualitative due to the variation in designs and measures. The next step is not simply a larger review. The field requires research based on authentic media content, which measures attention and load simultaneously, and which tests the listeners' movement between sound layers in time.

REFERENCES

- Blauenfeldt, F. B. (2025). *Designing a dynamic music system for narrative audio storytellers* (Master's thesis). IT University of Copenhagen, Denmark.
- Brockhoff, L., Vetter, L., Bruchmann, M., Schindler, S., Moeck, R., & Straube, T. (2023). The effects of visual working memory load on detection and neural processing of task-unrelated auditory stimuli. *Scientific Reports*, 13, Article 4342. doi:10.1038/s41598-023-31132-7
- Brown, V. A., & Strand, J. F. (2019). About face: Seeing the talker improves spoken word recognition but increases listening effort. *Journal of Cognition*, 2(1), 44. doi:10.5334/joc.89
- Calculus, A. (2024). Development of auditory scene analysis: A mini-review. *Frontiers in Human Neuroscience*, 18, Article 1352247. doi:10.3389/fnhum.2024.1352247
- Engstrom, E., & Stephenson, D. (2024). Audio and video journalism. In *Convergent journalism: An introduction* (pp. 99-119). London, England: Routledge.
- Ferrario, A., & Rankin, J. (2021). Auditory streaming emerges from fast excitation and slow delayed inhibition. *Journal of Mathematical Neuroscience*, 11, Article 8. doi:10.1186/s13408-021-00106-2
- Fintor, E., Aspöck, L., Fels, J., & Schlittmeier, S. J. (2021). The role of spatial separation of two talkers' auditory stimuli in the listener's memory of running speech: Listening effort in a non-noisy conversational setting. *International Journal of Audiology*, 61(5), 371-379. doi:10.1080/14992027.2021.1922765
- Gnanateja, G. N., Devaraju, D. S., Heyne, M., Quique, Y. M., Sitek, K. R., Tardif, M. C., ... Dial, H. R. (2022). On the role of neural oscillations across timescales in speech and music processing. *Frontiers in Computational Neuroscience*, 16, Article 872093. doi:10.3389/fncom.2022.872093
- Gurariy, G., Randall, R., & Greenberg, A. S. (2021). Manipulation of low-level features modulates grouping strength of auditory objects. *Psychological Research*, 85(6), 2256-2270. doi:10.1007/s00426-020-01391-4
- Gustafson, S. J., Nelson, L., & Silcox, J. W. (2023). Effect of auditory distractors on speech recognition and listening effort. *Ear and Hearing*, 44(5), 1121-1132. doi:10.1097/aud.0000000000001356
- Hermes, D. J. (2023). Auditory-stream formation. In *The perceptual structure of sound* (pp. 559-784). Cham, Switzerland: Springer International Publishing.
- Jhang, G.-Y., Ueda, K., Takeichi, H., Remijn, G. B., & Hasuo, E. (2025). Rivalry between pitch and timbre in auditory stream segregation. *PLOS ONE*, 20(6), Article e0323964. doi:10.1371/journal.pone.0323964
- Kankam, H. (2025). Perception and attention. In *A brief excursion into human cognition: The evolving influence of social media & artificial intelligence* (pp. 17-31). Cham, Switzerland: Springer Nature Switzerland.
- Kurzyński, M. (2025). Poly-temporal, multi-layered: A techno-cognitive theory of narrative experience in literature. *International Journal of Humanities and Arts Computing*, 19(1), 33-48. doi.org/10.3366/ijhac.2025.0343
- Liu, M., Arseneau-Bruneau, I., Farrés Franch, M., Latorre, M.-E., Samuels, J., Issa, E., ... Zatorre, R. J. (2025). Auditory working memory mechanisms mediating the relationship between musicianship and auditory stream segregation. *Frontiers in Psychology*, 16, Article 1538511. doi:10.3389/fpsyg.2025.1538511
- Mobin, S. (2025). *Interpreting nonverbal cues through sound: Exploring accessibility for BLV individuals in online communication using cinematic media and sound design* (Master's thesis). OCAD University, Toronto, Canada.
- Noyce, A. L., Kwasa, J. A. C., & Shinn-Cunningham, B. G. (2023). Defining attention from an auditory perspective. *WIREs Cognitive Science*, 14(1), Article e1610. doi:10.1002/wcs.1610
- Oh, Y., Zuwala, J. C., Salvagno, C. M., & Tilbrook, G. A. (2022). The impact of pitch and timbre cues on auditory grouping and stream segregation. *Frontiers in Neuroscience*, 15, Article 725093. doi:10.3389/fnins.2021.725093
- Que, Y., & Hu, X. (2025). Enhancing learners' reading comprehension with preferred background music: An eye-tracking, EEG, and heart rate study. *Reading Research Quarterly*, 60(2), Article e70004. doi:10.1002/rq.70004
- Que, Y., Zheng, Y., Hsiao, J. H., & Hu, X. (2025). Using eye movements, electrodermal activities, and heart rates to predict different types of cognitive load during reading with background music. *Scientific Reports*, 15, Article 32635. doi:10.1038/s41598-025-03052-1
- Ragert, M., Fairhurst, M. T., & Keller, P. E. (2014). Segregation and integration of auditory streams when listening to multi-part music. *PLoS ONE*, 9(1), Article e84085. doi:10.1371/journal.pone.0084085
- Rennig, J., Wegner-Clemens, K., & Beauchamp, M. S. (2020). Face viewing behavior predicts multisensory gain during speech perception. *Psychonomic Bulletin & Review*, 27(1), 70-77. doi:10.3758/s13423-019-01665-y

- Roberts, B., & Haywood, N. R. (2023). Asymmetric effects of sudden changes in timbre on auditory stream segregation. *Journal of the Acoustical Society of America*, 154(1), 363-378. doi:10.1121/10.0020172
- Symonds, R. M., Zhou, J. W., Cole, S. L., Brace, K. M., & Sussman, E. S. (2019). Cognitive resources are distributed among the entire auditory landscape in auditory scene analysis. *Psychophysiology*, 57(4), Article e13487. doi:10.1111/psyp.13487
- Squires, V. (2023). Thematic analysis. In *Varieties of qualitative research methods: Selected contextual perspectives* (pp. 463-468). Cham: Springer International Publishing.
- Thomassen, S., Hartung, K., Einhäuser, W., & Bendixen, A. (2022). Low-high-low or high-low-high? Pattern effects on sequential auditory scene analysis. *Journal of the Acoustical Society of America*, 152(5), 2758-2768. doi:10.1121/10.0015054
- Vosburg, M. (2024). Watching the watchers: A meta-thematic analysis of media effects studies of visual communication. *Review of Communication Research*, 12, 67-85. doi:10.52152/RCR.V12.5
- Yang, Y., Sannusi, S. N., & Ahmad Rizal, A. R. (2025). The impact of gratifications on the adoption of immersive journalism: A meta-analysis. *Review of Communication Research*, 13, 81-95. doi:10.52152/RCR.V13.7