

Review

Nonverbal Auditory Communication for Human–Robot Interaction in Industry 5.0: A Scoping Review

Tom Schmid ^{1,2,*} , Manja Lohse ^{1,2}, Sven Winkelmann ^{1,3}  and Alexander von Hoffmann ^{1,2,3} 

¹ Elektrotechnik Feinwerktechnik Informationstechnik, Technische Hochschule Nürnberg Georg Simon Ohm, Keßlerplatz 12, 90489 Nuremberg, Germany; manja.lohse@th-nuernberg.de (M.L.); sven.winkelmann@th-nuernberg.de (S.W.); alexander.vonhoffmann@th-nuernberg.de (A.v.H.)

² Technologietransferzentrum (TTZ) Nürnberger Land, Martin-Luther-Straße 18, 91207 Lauf an der Pegnitz, Germany

³ Ohm User Experience Center (Ohm-UX), Brucknerstraße 11, 90429 Nuremberg, Germany

* Correspondence: tom.schmid@th-nuernberg.de

Abstract

In Industry 5.0 (I5.0), close-proximity human–robot collaboration demands communication beyond conventional alarms and speech. Nonverbal auditory communication offers a complementary modality, yet its role in I5.0 remains unmapped. This scoping review maps nonverbal auditory communication research in I5.0 Human–Robot Interaction (HRI) and compares it with general HRI literature to identify transfer potential and research gaps. Peer-reviewed English-language articles (2023–April 2026) addressing nonverbal sound in HRI contexts were included. Speech, emotion detection, haptic interfaces and non-HRI domains were excluded. A search with two syntaxes across Web of Science, Scopus, IEEE Xplore, ACM and MDPI, supplemented by citation searching, targeted I5.0-specific (Syntax S1) and general HRI auditory literature (Syntax S2). This created two article record sets, n1 and n2. Articles were organized following Arksey and O'Malley's framework and PRISMA-ScR into four inductively derived clusters: Sonification, Multimodal Feedback Systems, Safety and Frameworks and Concepts. From 782 initial records, 16 (n1) and 32 (n2) articles were included. In I5.0, multimodal feedback dominates: intentionally designed nonverbal sounds improve situational awareness, reduce cognitive workload and increase perceived safety. Compared to n2, which is shaped by social robotics and emotion-driven sound design, five gaps emerge in I5.0: absent emotion-related sound perception research, missing field studies, missing industry-specific sound design frameworks, underutilized sonification for spatial awareness and safety and no unimodal auditory studies under realistic industrial conditions. A dedicated sound design framework operationalizing I5.0 communicative requirements into designable sound parameters is needed, alongside empirical validation under realistic industrial noise conditions.

Keywords: human-robot interaction; industry 5.0; nonverbal auditory communication; nonverbal sound; robotics; human–robot collaboration



Academic Editor: Stefano Baraldo

Received: 29 May 2026

Revised: 22 June 2026

Accepted: 23 June 2026

Published: 26 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the emergence of human-centered Industry 5.0 (I5.0) [1], close-proximity human–robot interaction (HRI) has become a central concept of industrial manufacturing. Unlike Industry 4.0 (I4.0), which prioritized automation and reduced human involvement, I5.0 requires operators to perceive and interpret robot behavior dynamically. Effective communication between humans and robots is therefore a core requirement. Since no

single modality can reliably cover the full range of informational demands, this calls for multimodal HRI [2,3].

Multimodal HRI research has grown in recent years utilizing visual, haptic and auditory modalities [2]. Within the auditory domain, speech is regarded as the most intuitive modality [4] and is therefore one of the most studied [3,5–7]. Nevertheless, speech faces intrinsic limitations in industrial environments: speech intelligibility depends fundamentally on signal-to-noise ratio, recognition degrades under environmental noise [2,8,9] and spoken interaction can increase the cognitive load of workers [10–12].

These limitations motivate exploring nonverbal auditory communication as a complementary modality. Unlike speech, nonverbal sound is independent of linguistic intelligibility. Drawing from sonification research, nonverbal sounds may support continuous awareness of robot motion, state and intention in ways that speech cannot easily provide [13,14].

Nonverbal sound in HRI refers to the auditory signals produced by robots, including auditory icons, earcons or audio cues, each reflecting different design principles. Even sounds that emerge from the robot's movement (consequential sounds) are considered part of this category. This terminological fragmentation has been addressed by Zhang and Fitter [14], who consolidated a taxonomy with the generalized term of nonverbal HRI sounds.

Speech and nonverbal sound, however, operate within shared auditory constraints that industrial environments impose. Environmental noise occupies and degrades the auditory channel [11,12], additional signals compete for limited attentional resources [10] and sounds risk being masked by background noise [15]. Norms such as ISO 7731 [16], DIN 33404-3 [17], ISO 8201 [18] and ISO 11429 [19] address these constraints by specifying design requirements optimized for detection through loudness, spectrum and rhythm.

While these norms have shaped industrial sounds around warnings and notifications, broader HRI research suggests that nonverbal sound can be used for communicative functions [14]. Robinson et al. [20] argue that both the generation and perception of nonverbal sound are essential for effective HRI. However, these approaches have emerged in social and general HRI, which differ structurally from I5.0, leaving the potential transferability and the current state of nonverbal sound research in I5.0 unknown.

While Zhang and Fitter [14] provide a systematic review of nonverbal HRI sound up to 2023, their work does not address I5.0 or examine whether nonverbal HRI sound research has shifted toward industrial applications. As I5.0 is still emerging, no review has yet mapped the current evidence at this intersection or identified concrete research gaps. This paper addresses both aspects through a scoping review at the intersection of nonverbal auditory communication and I5.0, motivating the following research questions (RQ):

- RQ1: Which nonverbal sound applications are reported in the literature for HRI in I5.0 contexts and what key findings related to their use are identified?
- RQ2: How does the representation of nonverbal sounds in I5.0 HRI compare to its application in broader HRI contexts and what gaps can be identified?

Answering RQ1 provides the first structured mapping of nonverbal sound research explicitly situated in I5.0. Comparing this with the broader HRI literature (RQ2) shows where established approaches already exist and simply have not yet been transferred to industrial settings. Through this study, research gaps emerge that give future work a starting point for advancing nonverbal auditory communication in I5.0 systematically.

2. Methodology

The research questions focus on analyzing current applications and key findings at the intersection of nonverbal auditory HRI and I5.0 manufacturing. As I5.0 was introduced

in 2021 [1] and the evidence base remains limited, a scoping review is appropriate for mapping emerging literature [21]. The study followed the methodology of Arksey and O'Malley's [21] scoping review framework. Additionally, the PRISMA Extension for Scoping Reviews checklist [22] was applied (documented in Supplementary Materials S1 (Sup1_PRISMA_ScR_Checklist.pdf)). No review protocol was registered or published for this study. The search and selection process is illustrated using the PRISMA 2020 flow diagram [23]. Data charting was conducted by one reviewer and discussed between the four authors to ensure consistency and to reduce potential selection bias in the screening process. As a result, the overall phases of the scoping review are: Identification, Screening and Inclusion (Figure 1).

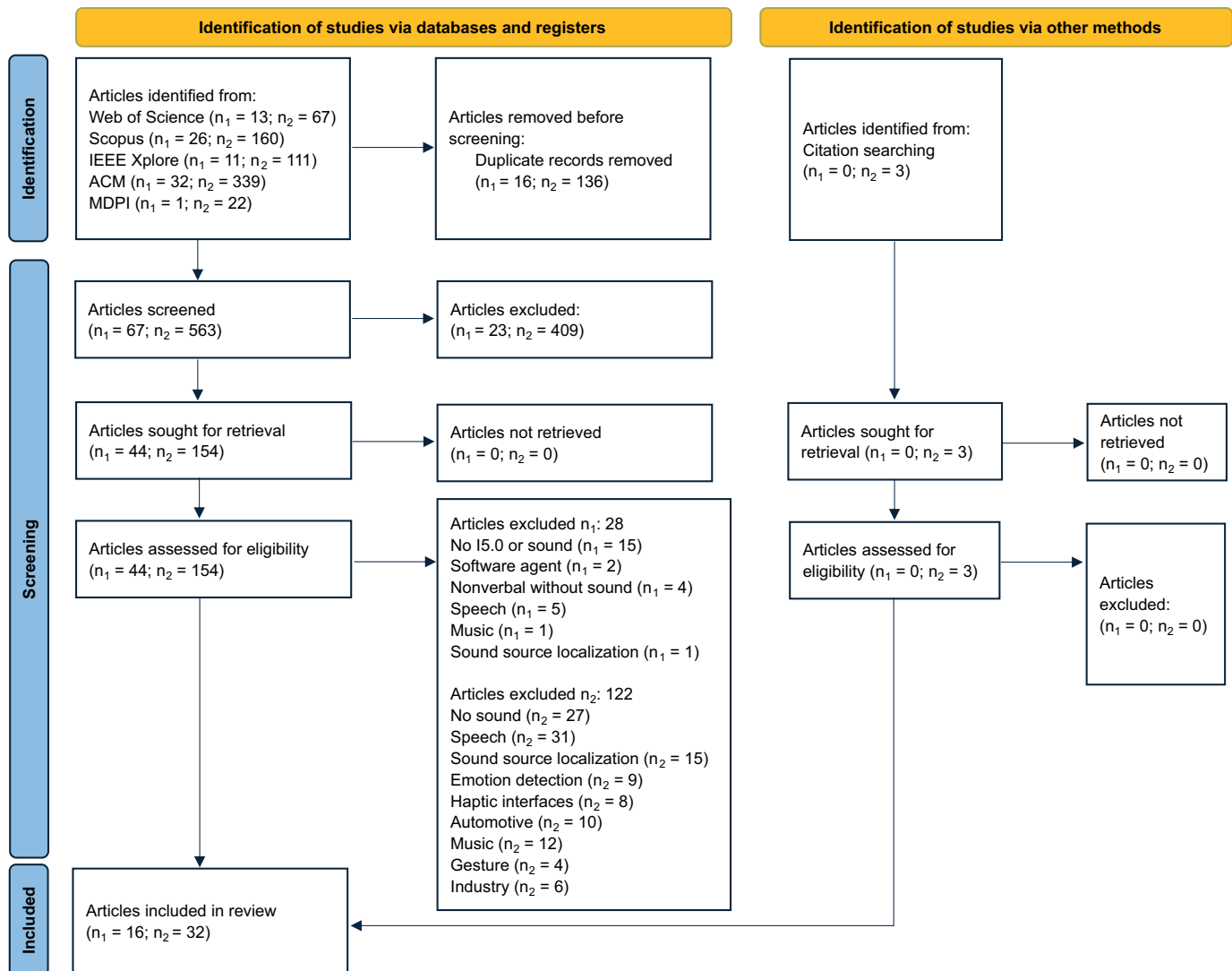


Figure 1. Scoping review flow diagram based on PRISMA statement 2020 [23].

2.1. Identification

The identification period of this review ranged from 2023 to 18 April 2026. This period was chosen because I5.0 was conceptually established in 2021, making earlier work mostly irrelevant. Moreover, Zhang and Fitter's 2023 systematic review confirms this assumption as it already provides coverage of nonverbal HRI sound research up to 2023 and does not mention I5.0. Due to the comprehensive coverage of HRI research up to 2023 provided by Zhang and Fitter [14], articles published between 2021 and 2022 are indirectly represented through their review. Nevertheless, it should be noted that relevant

manufacturing-focused sonification studies, particularly those not explicitly situated within HRI, may have been overlooked.

In the identification phase, queries were defined and iteratively refined based on preliminary searches to balance recall and result volume. Terms on sound related topics originated from the taxonomy of Zhang and Fitter [14]. To answer the proposed research questions RQ1 and RQ2, two specific query syntaxes Syntax S_1 and Syntax S_2 were constructed. The full list of syntaxes is documented in the Supplementary Materials S2 (Sup2_Syntaxes_Fulllist.xlsx). As an example the syntaxes for the database Scopus are presented:

- S_1 : TITLE-ABS-KEY (("human-robot interaction" OR "human-robot collaboration" OR "human-robot cooperation" OR "human-robot communication" OR HRI OR HRC) AND ("Industry 5.0" OR "smart manufactur*" OR "collaborative assembl*" OR "collaborative manufactur*" OR "human-centric manufactur*" OR "smart production" OR "resilient manufactur*" OR "sustainable manufactur*" OR "symbiotic manufactur*" OR "human-robot coexistence" OR "operator 5.0" OR "collaborative production" OR cobot* OR "collaborative robot*" OR "autonomous mobile robot*" OR AMR) AND (auditory* OR "nonverbal sound*" OR "robot sound*" OR earcon* OR "acoustic feedback" OR "acoustic cue*" OR "acoustic alert*" OR sonification OR "sound design" OR soundscape* OR AVAS OR "consequential sound*" OR "functional sound*")) AND PUBYEAR > 2022 AND (DOCTYPE (ar) OR DOCTYPE (cp) OR DOCTYPE (re))
- S_2 : TITLE-ABS-KEY (("human-robot interaction" OR "human-robot collaboration" OR "human-robot cooperation" OR "human-robot communication" OR HRI OR HRC) AND (auditory* OR "nonverbal sound*" OR "robot sound*" OR earcon* OR "acoustic feedback" OR "acoustic cue*" OR "acoustic alert*" OR sonification OR "sound design" OR soundscape* OR AVAS OR "consequential sound*" OR "functional sound*")) AND PUBYEAR > 2022 AND (DOCTYPE (ar) OR DOCTYPE (cp) OR DOCTYPE (re))

The used databases were Web of Science (WoS), Scopus, Institute of Electrical and Electronics Engineers (IEEE) Xplore, Association for Computing Machinery Digital Library (ACM), and Multidisciplinary Digital Publishing Institute (MDPI). Query syntax was adapted for each database, affecting only the format, not the content. The initial search in each database focused on titles, abstracts and keywords. To achieve this, the following search fields were used for each database:

- WoS: Topic—includes title, abstract, Keywords Plus and author keywords;
- Scopus: TITLE-ABS-KEY—includes title, abstract and keywords;
- IEEE Xplore: All Metadata—includes abstract, index terms and bibliographic citation data;
- ACM: Anywhere (AllField)—this option was used to keep the query syntax consistent. However, it may have introduced additional noise in the form of false-positive articles, as it searches beyond titles, abstracts and keywords;
- MDPI: Title, Abstract, Keywords.

In addition, citation searching was applied. This review focused exclusively on English-language and peer-reviewed articles. This criterion was already ensured through the database search filters. The approach with two syntaxes resulted in two record sets of articles, n_1 and n_2 . Due to the two-syntax approach across multiple databases, duplicates were removed within each record set.

2.2. Screening

The identified articles were screened on titles and abstracts to enable first-stage exclusion. Therefore, the exclusion criteria were iteratively refined and applied to both record sets to better fit their respective objectives. For n_1 , articles were excluded if an I5.0 or

sound-related context was absent, whereas for n_2 the only exclusion criterion was the absence of a sound-related context.

Eligibility was assessed through full-text review. Each article was rated according to its topical relevance to the corresponding RQ and the prominence of sound as “relevant”, “discussion” or “not relevant”. Articles rated with “discussion” were discussed in the research group to minimize bias. With this rating process, exclusion criteria for n_1 and n_2 were developed separately and are documented in Figure 1 and Section 3. In line with scoping review methodology [21], no formal quality appraisal of included articles was conducted. It is important to note that the reviewed articles may focus on different primary aspects of HRI, with nonverbal sounds appearing as a secondary or supplementary element rather than the central focus. The complete review process, including the exclusion criteria and the number of articles remaining at each step, is shown in Figure 1 and Section 3.

2.3. Inclusion

Thematic clusters were formed inductively during full-text assessment of n_1 and iteratively refined. To enable the direct comparison required by RQ2, the articles in n_2 were subsequently mapped onto the thematic cluster structure. While this approach may obscure patterns unique to n_2 , the focus is intentional as the goal was to assess what the broader HRI literature offers relative to I5.0 requirements. Findings within each cluster were synthesized narratively by summarizing and interpreting the applications and key findings across included articles. For further analysis, bibliographic metadata (author, year, title) and the content dimensions “Application” and “Key Finding” related to nonverbal auditory communication were listed within each cluster. Due to the relatively small number of included articles, sound-related key findings were extracted manually by reviewing the conclusions and results of each article.

3. Results

The identification was conducted on 18 April 2026 and Figure 1 shows the review process with numbers of included and excluded articles of each phase. The two record sets initially yielded $n_{init} = 782$ articles in all databases. After duplicate removal, $n_1 = 67$ and $n_2 = 563$ articles remained (Figure 1).

During title and abstract screening, the number of records was reduced based on the exclusion criteria explained in Section 2.2 and additionally identified criteria during the process. For n_1 , the topics medicine and agriculture, as well as the technologies drones and exoskeletons were excluded as they do not fit the I5.0 paradigm. The exclusion criterion for n_2 was the absence of sound context. Articles were removed if no sound-related keyword from the query appeared in the title or abstract. This resulted in $n_1 = 44$; $n_2 = 154$ articles.

In the eligibility phase, articles were excluded from n_1 if I5.0 or the sound context were only marginally mentioned. This resulted in the exclusion of the following topics: software agents, nonverbal communication without sound, speech, music and sound source localization. Research domains such as digital twins and virtual reality used for prototyping were included if they involved a robotic system as these environments serve as a relevant testbed for HRI. Since n_2 focused on general HRI, the exclusion criteria were less strict research domain-wise and more focused on topics. The identified exclusion criteria were speech, sound source localization, emotion detection, haptic interfaces, automotive, music, and nonverbal communication with gesture and industry and were therefore removed.

Although cobots and autonomous mobile robots (AMR) are elements of I5.0 and part of the I5.0 query, they were not strictly excluded from n_2 because they also appear independently of industry. Additional citation searching contributed $S_{CS} = 3$ more articles to n_2 , resulting in a final number of $n_{S1} = 16$ and $n_{S2} = 32$, presented by publication year

in Figure 2. As shown in Figure 2, the overall number of publications in this research field is rising, consistent with the volume reported by Zhang and Fitter [14]. Publications related to I5.0 increased in 2024, suggesting a growing interest in this field. However, a definitive trend cannot yet be established, as no publications from 2026 met the eligibility criteria for inclusion in this review.

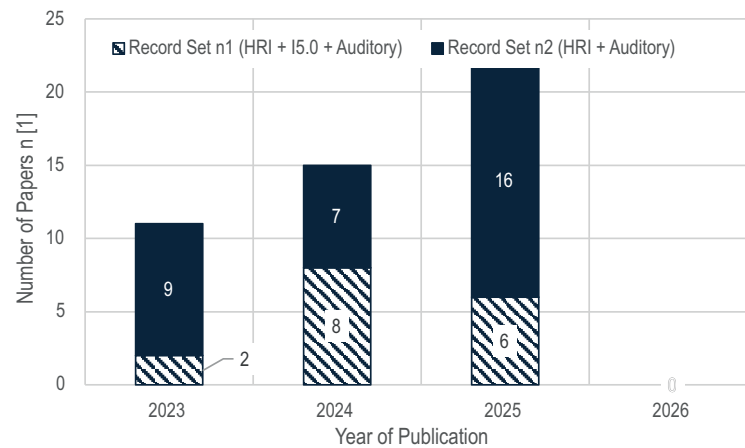


Figure 2. Relevant articles of record set n_1 and n_2 ; articles are only included up to 18 April 2026, so no relevant articles have been published yet but may appear later in 2026.

Based on the 16 articles of n_1 , four clusters were identified, representing the core thematic structure of the I5.0 articles:

1. Sonification;
2. Multimodal Feedback System;
3. Safety;
4. Frameworks and Concepts.

In the following sections, the individual clusters are presented sequentially, with a summary followed by a table. The articles are organized alphabetically by year in the clusters (Tables 1–8). Tables 1–4 cover n_1 and are discussed in Section 3.1. Tables 5–8 cover n_2 and are discussed in Section 3.2.

3.1. Nonverbal Sound Applications in Industry 5.0 (Syntax 1)

Two studies (Table 1) place sonification at the center of their research within I5.0. Kacem et al. [24] and Roy et al. [25] address sonification of robot behavior: Kacem et al. map AMR motion to nonverbal sound via Wave Space Sonification, while Roy et al. modulate acoustic parameters such as pitch and rhythm through reinforcement learning to convey functional robot states.

Table 1. Record set n_1 : cluster 1—Sonification.

Year	Authors	Title	Application	Key Finding
2024	Kacem et al. [24]	Transforming Motion into Sound: A Novel Sonification Approach for Teams of Mobile Robots	Wave Space Sonification for AMR movement	Multi-robot motion can be sonified using Wave Space Sonification
2024	Roy et al. [25]	Learning to Communicate Functional States With Nonverbal Expressions for Improved Human-Robot Collaboration	State communication in human–robot collaboration	Personalized nonverbal sounds improve humans’ accuracy in identifying a robot’s functional state

Seven articles focused on robots' multimodal output. Therefore, the cluster multimodal feedback system (Table 2) was introduced. D'Attanasio et al. [26] and Huang et al. [27] examined the role of nonverbal sounds within such system through review-based contributions, while other studies [28–30] present empirical or framework-based implementations. Both Dei et al. [28] and Mizanoor Rahman [29] report that performance and cognition in collaborative tasks can be improved by combining visual and auditory modalities. Javernik et al. [31] find that robot sound level alone did not significantly affect worker efficiency. Grego et al. [30] present nonverbal sound as a supportive modality within an augmented reality framework for industrial digital twins.

Table 2. Record set n_1 : cluster 2—Multimodal Feedback System.

Year	Authors	Title	Application	Key Finding
2023	Javernik et al. [31]	Audio-Visual Effects of a Collaborative Robot on Worker Efficiency	Collaborative Assembly	Robot sound level alone does not significantly affect worker efficiency
2024	D'Attanasio et al. [26]	Exploring User Preferences in Multimodal Human-Robot Interaction for the Human-Centered Design of an Assistant Robot	Human-centered design of multimodal human–robot interaction with Cobots	Nonverbal sound is important for communicating robot actions; Multimodality preferred
2024	Dei et al. [28]	Design and testing of (A)MICO: a multimodal feedback system to facilitate the interaction between cobot and human operator	Multimodal feedback for tasks	Nonverbal sound, when tightly coupled with visual cues, significantly improves users' understanding of cobot states and errors
2024	Grego et al. [30]	Enhancing Human-Machine Interactions: A Novel Framework for AR-Based Digital Twin Systems in Industrial Environments	Nonverbal sound is implemented as a supportive modality	More research suggested on multimodal feedback including sound
2024	Huang et al. [27]	Towards instructional collaborative robots: From video-based learning to feedback-adapted instruction	Feedback-adaptive instructional cobots for human-centric manufacturing	Auditory verbal instructions and sound signals are essential for effective robot-guided teaching, especially when visual channels are overloaded
2025	Mizanoor Rahman [29]	Communications in a Trust-Triggered Cyber-Physical-Human System for Human-Robot Collaborative Manufacturing	Trust-aware human–robot collaborative assembly	Augmenting trust communication with nonverbal sound improves performance and cognition

The four articles in Cluster 3 (Table 3) share a common focus on safety-relevant auditory communication in human–robot workspaces. While Haney and Liang [32] identify nonverbal sounds as an insufficiently studied factor in perceived safety through a literature review, the remaining three studies [33–35] provide empirical evidence. Kamezaki et al. [33] and San Martin et al. [34] report improvements in perceived safety when nonverbal sound is presented alongside visuals. Hostettler et al. [35] report that non-adaptive consequential robot sounds were perceived as stressful, while adaptive robot behavior reduced discomfort.

Table 3. Record set n_1 : cluster 3—Safety.

Year	Authors	Title	Application	Key Finding
2024	Haney and Liang [32]	A Literature Review on Safety Perception and Trust during Human-Robot Interaction with Autonomous Mobile Robots That Apply to Industrial Environments	Safer AMRs in manufacturing and warehousing.	Auditory cues in AMRs are promising but insufficiently studied, with mixed evidence on improving perceived safety and intent understanding
2024	Kamezaki et al. [33]	Dynamic Collaborative Workspace Based on Human Interference Estimation for Safe and Productive Human-Robot Collaboration	Safe and productive parallel work	Visual and auditory state indicator improve safety and predictability
2025	Hostettler et al. [35]	Real-Time Adaptive Industrial Robots: Improving Safety And Comfort In Human-Robot Collaboration	Adaptive collaboration for Safety and Comfort with Cobot	Non-adaptive consequential sounds were perceived as stressful, while adaptive robot behavior reduced discomfort
2025	San Martin et al. [34]	Mixed reality representation of hazard zones while collaborating with a robot: sense of control over own safety	Nonverbal sounds for communicating safety	Warnings using sound and visual improve feelings of safety and control without reducing performance

Cluster 4 (Table 4) contains four conceptual and framework-oriented contributions on auditory HRI in I5.0 contexts. Rather than reporting isolated empirical results, these contributions position nonverbal sound within broader topics. Chu et al. [36] and Schneiders et al. [37] treat nonverbal sound as a general communication tool of collaborative systems. Chu et al. [36] use sound in both directions, as an input modality for human intention recognition and as an output for warning sounds. In the Schneiders et al. [37] study, participants used consequential sounds unintentionally enhancing task performance. This inspired design recommendations. Zhang et al. [38] address the design process itself, reporting that effective robot sound design requires interdisciplinary collaboration between roboticists and sound designers. Wang et al. [39] take a broader perspective through a systematic review of XR-enabled remote HRI, identifying multimodal feedback as promising for improving perception of remote environments.

Table 4. Record set n_1 : cluster 4—Frameworks and Concepts.

Year	Authors	Title	Application	Key Finding
2023	Zhang et al. [38]	Hearing it Out: Guiding Robot Sound Design through Design Thinking	Collaborative design guidelines	Robot sound design requires an interdisciplinary development process
2025	Chu et al. [36]	A Game-Theoretic and Multimodal Interaction Framework for Collaborative Robots in Smart Manufacturing	Multimodal sensing (including sound)	Nonverbal sound serves as an effective input modality for cobots to support human intention recognition
2025	Schneiders et al. [37]	Understanding Entrainment in Human Groups: Optimising Human-Robot Collaboration from Lessons Learned during Human-Human Collaboration	HRI for Industry	Nonverbal sound is important for state communication. Task noise is used for feedback
2025	Wang et al. [39]	A Systematic Review of XR-Enabled Remote Human-Robot Interaction Systems	XR-enabled robot teleoperation systems	Nonverbal sound and multimodal feedback are promising for improving perception of remote environment

3.2. Nonverbal Sound in HRI (Syntax 2)

For n_2 , the same cluster structure is applied, maintaining comparability across both record sets. The Sonification cluster contains 16 articles with a partially different application set compared to n_1 . Cluster 1 of n_2 (Table 5) includes studies on emotional non-linguistic utterances and on sound parameters such as pitch and timbre map to specific emotions [40–44]. A second segment covers movement sonification [45–49]. Further contributions examine consequential and ambient sound design for specific robot platforms such as delivery robots and spatial sound installations [50,51]. The cluster also includes semiotic and film-inspired analyses of robot sound as a communication medium [45], as well as works on auditory display classification systems that define how multiple robot states can be simultaneously encoded in sound [52,53].

Table 5. Record set n_2 : cluster 1—Sonification.

Year	Authors	Title	Application	Key Finding
2023	Latupeirissa and Bresin [54]	PepperOSC: enabling interactive sonification of a robot's expressive movement	Interactive sonification of robot movement	Movement-based sonification enhances expressiveness and aesthetic quality of robot actions
2023	Latupeirissa et al. [46]	Probing Aesthetics Strategies for Robot Sound: Complexity and Materiality in Movement Sonification	Sonification of robot movement	Aesthetically designed sounds are preferred; In real-world contexts, subtle sounds are preferred.
2023	Liu et al. [42]	Robots' "Woohoo" and "Argh" Can Enhance Users' Emotional and Social Perceptions: An Exploratory Study on Non-lexical Vocalizations and Non-linguistic Sounds	Emotion in social robotic sounds with natural human non-lexical vocalizations	Significantly improves emotion recognition, trust and perceived naturalness compared to synthesized, musical or no sounds
2023	Mannone et al. [49]	The Sound of Swarm. Auditory Description of Swarm Robotic Movements	Movement sonification for swarm robotics	Mapping motion parameters to musical features enables listeners to perceive emergent swarm dynamics
2023	Orthmann et al. [53]	Sounding Robots: Design and Evaluation of Auditory Displays for Unintentional Human-robot Interaction	HRI with mobile robotics; layer-based sonification	Possibility to communicate multiple aspects of a robot's state through sound
2023	Robinson et al. [51]	Spatially Distributed Robot Sound: A Case Study	Spatial audio with robots and additional speakers	Spatially distributed robot sound changes how people engage with robots
2024	Andrade Pires et al. [43]	Emotive Acoustics: Sound Design in Robotic Emotion Expression. A Study on Participant Generated Sounds	Emotion in robotic sounds	Significance of emphasizing human-like sounds, with a focus on onomatopoeia
2024	Cuan et al. [48]	Music Mode: Transforming Robot Movement into Music Increases Likability and Perceived Intelligence	Motion sonification with instruments	Participants perceived the robot as safer, more intelligent and more likable accompanied by music-based sonification, compared to mechanical or no sonification

Table 5. *Cont.*

Year	Authors	Title	Application	Key Finding
2024	Khota et al. [40]	Synthesis of Non-Linguistic Utterances for Sound Design Support Using a Genetic Algorithm	Automatic sound design for social robots	Genetically generated nonverbal sounds can convey emotion
2024	Wolfe et al. [44]	Dimensional Design of Emotive Sounds for Robots	Emotion sound design for robots with non-linguistic utterances	Mapping musical/prosodic parameters onto a valence–arousal model enables robots to convey excitement and sadness
2025	Abboud and Bruno [41]	Automated Generation of Emotion-Conveying Sounds for Robots	Emotion classification of robotic sounds	Emotion recognition approaches from speech/music do not transfer well to robotic sounds; New approach positions emotions in the Valence–Arousal space more reliable
2025	Arreghini et al. [52]	Hearing the Robot’s Mind: Sonification for Explicit Feedback in Human-Robot Interaction	Feedback sonification (intention prediction) with sound synthesis	Sonification enhances communication of robot states
2025	Latupeirissa et al. [45]	Semiotic Analysis of Robot Sounds in Films: Implications for Sound Design in Social Robotics	Deriving sound design guidance for social robots from film	Robot sounds in film use pitch and timing to convey identity, emotion and narrative meaning
2025	Naphausen et al. [47]	New Design Potentials of Non-mimetic Sonification in Human-Robot Interaction	Non-mimetic Sonification	Mapping robot motion and internal state to abstract sound parameters supports awareness and interaction
2025	Zhang et al. [50]	A Case Study on Robot Sound Design for a Sidewalk Delivery Robot	Real-world sidewalk delivery robots	Carefully designed nonverbal robot sounds increase perceived social interaction

The Multimodal Feedback Systems cluster (Table 6) includes four articles. Two articles address teleoperation and supervisory control contexts, where multimodal communication reduces cognitive load and improves situational awareness [55,56]. Clinical and assistive applications are represented, including robotic ultrasound with auditory and visual feedback [57]. The cluster further includes a social and affective system where haptic and auditory stimuli are combined to support anxiety reduction [58].

The Safety cluster (Table 7) focuses on how robot sounds function as spatial and safety-relevant signals in shared environments. One segment covers consequential sounds and how these affect human perception of robots [59,60]. Closely related are studies on auditory detectability and localization, examining how well humans can detect and locate robots by sound alone [15,61]. Navigation assistance for blind and low-vision users, where auditory cues serve as safety-critical orientation signals, is represented in [62,63]. One contribution extends safety to extreme environments by using sonification of radiation data to protect teleoperation operators [64].

Table 6. Record set n_2 : cluster 2—Multimodal Feedback System.

Year	Authors	Title	Application	Key Finding
2024	Attfield et al. [55]	A comparison of audible, visual, and multi-modal communication for multi-robot supervision and situational awareness	Nonverbal communication methods in HRI	Multimodal auditory and visual method outperformed both the purely visual and auditory-based methods in performance and situation awareness
2025	Matheus et al. [58]	Ommie: The Design and Development of a Social Robot for Anxiety Reduction	Mental health support via socially assistive robotics	Haptic interaction and nonverbal sounds can reduce user anxiety
2025	Song et al. [57]	Exploring Multisensory Feedback for Robotic Ultrasound in XR	Medical robotics with cobots	Multimodal feedback enhances user experience
2025	Yang et al. [56]	A Human-Factor-Driven Evaluation of Multi-Modal Feedback for Enhancing Teleoperated Microrobotic Swarm Control	Teleoperated microrobotic swarm control in biomedical tasks	Adding nonverbal sound to haptic–visual interfaces significantly improves swarm stability and reduces operator cognitive load

Table 7. Record set n_2 : cluster 3—Safety.

Year	Authors	Title	Application	Key Finding
2023	West et al. [64]	Sonification of Ionising Radiation Data for Robot Operators	Reducing visual cues in extreme environments for mobile robots	Realistic stochastic Geiger-counter-like sounds that sonify radiation rate provide effective auditory communication for robot operators
2024	Allen et al. [59]	Sound Judgment: Properties of Consequential Sounds Affecting Human-Perception of Robots	Consequential sounds of mobile robots	Participants expressed a clear preference for informative and rhythmic sounds over loud or high-pitched noises
2024	Maehigashi et al. [65]	Impacts of Robot Beep Timings on Trust Dynamics in Human-Robot Interaction	Trust repair in human–robot interaction	Simple robot beep sounds can modulate trust dynamics as strongly as verbal cues when timed around successful actions, but are ineffective for trust repair after errors compared to verbal apologies
2025	Agrawal et al. [15]	Sound Matters: Auditory Detectability of Mobile Robots	Safe and intuitive co-presence of humans and mobile robots when robots operate outside visual field	Motion sounds from a quadruped robot are detected far earlier than continuous wheeled robot noise, even at equal loudness and in noisy environments
2025	Allen et al. [60]	Robots Have Been Seen and Not Heard: Effects of Consequential Sounds on Human-Perception of Robots	Consequential sounds of robots	Consequential sounds produced by robots significantly affect human perceptions leading to negative feelings

Table 7. *Cont.*

Year	Authors	Title	Application	Key Finding
2025	Currie et al. [66]	Sonic Sleight of Hand: Sound Induces Illusory Distortions in the Perception and Prediction of Robot Action	Perception and prediction of robot actions	Sounds produced by a robot during its movements significantly influence how observers perceive and predict the robot's actions
2025	Lee et al. [67]	Auditory and visual alarm designs impact clinicians' perceived cognitive workload	Interpretation of vital signs in healthcare	Adapting visual and auditory alarms reduces cognitive workload without performance loss
2025	Lee et al. [62]	Development of smart navigation robot for the visually impaired	Assistive indoor navigation for visually impaired	Real-time auditory alerts and verbal cues combined with AI-based object detection significantly improve spatial awareness
2025	Shih et al. [63]	A Multi-Method Investigation of Guide Robot Characteristics for Blind and Low-Vision User	Navigation of visually impaired with assistive robotics	Too much or poorly timed sound reduces comfort, agency and perceived safety
2025	Wessels et al. [61]	Auditory Localization and Assessment of Consequential Robot Sounds: A Multi-Method Study in Virtual Reality	Auditory localization of mobile robots	Auditory localization differs based on the robot type

The frameworks and concepts cluster (Table 8) includes systematic literature reviews covering auditory alerts in human–machine interfaces [68] and HRI for blind and low-vision users [69]. Conceptual contributions propose new taxonomies or communication codes, notably Urakami and Seaborn [70], who offer a communication-studies perspective on nonverbal HRI sounds across all five sensory modalities.

Table 8. Record set n_2 : cluster 4—Frameworks and Concepts.

Year	Authors	Title	Application	Key Finding
2023	Nees and Liebman [68]	Auditory Icons, Earcons, Spearcons, and Speech: A Systematic Review and Meta-Analysis of Brief Audio Alerts in Human-Machine Interfaces	Auditory communication in general	Speech, spearcons, and hybrid alerts generally outperform auditory icons and especially earcons on accuracy, reaction time and user preference, with similar workload and dual-task effects across types
2023	Urakami and Seaborn [70]	Nonverbal Cues in Human-Robot Interaction: A Communication Studies Perspective	Designing human–robot interaction by mapping human nonverbal communication	Nonverbal sounds such as paralanguage are central for conveying affect, engagement and empathy
2025	Wei et al. [69]	Human Robot Interaction for Blind and Low Vision People: A Systematic Literature Review	Assistive robotics for navigation of blind users	Auditory feedback (especially synthesized speech) is widely used for guidance and alerts

4. Discussion

Zhang and Fitter [14] provide the baseline for this work, reviewing articles on non-verbal HRI sound up to 2023 and introducing structured taxonomies. At the time of their review, published work concentrates mainly on social robots and controlled laboratory settings. Nevertheless, the proposed taxonomies of Zhang and Fitter [14] are also appropriate for I5.0 contexts and do not need adaptation. While Figure 2 indicates an increase toward industrial contexts in recent publications, this field remains less mature and more fragmented than the broader HRI sound literature. The discussion is based on key findings related to auditory communication.

4.1. Current Direction of Nonverbal Sounds in Industry 5.0

Following RQ1, the applications and key findings of Section 3.1 are interpreted below. In all clusters, nonverbal sounds are reported to contribute to HRI when intentionally designed and contextually integrated.

In n_1 , multimodal feedback systems represent the dominant research paradigm. Studies report that combining nonverbal sounds with visual and haptic signals reduces cognitive workload, and improves trust, situational awareness and worker efficiency in collaborative manufacturing tasks. These findings are consistent with Wickens' multiple resource theory [10]. In multimodal feedback systems [28–31] (Table 2), sounds are either deliberately designed or emerge from robot motion as consequential sounds [14], both of which can convey meaningful information in I5.0 contexts [14]. This difference carries direct design implications: deliberately designed sounds require an explicit sound design process, while consequential sounds demand awareness of unintended perceptual effects. Schneiders et al.'s [37] finding unintentionally confirms this within a pick-and-place task by showing that collaborators relied on sound produced by the task to time their actions and free visual attention. This suggests that consequential sounds can create implicit synchronization between operators and cobots.

A major limitation of existing studies on multimodal feedback systems is that the sound design process is rarely documented. This indicates that sound design choices are often made ad hoc, without considering user perception. Furthermore, Zhang et al. [38] show through two qualitative user studies that roboticists prioritize communicative function while sound designers focus on aesthetic form. This indicates the need for collaboration to develop interdisciplinary design frameworks.

Safety-related auditory communication receives comparatively little attention in multimodal feedback systems, and where it appears, sounds are predominantly used as attention anchors. While Kamezaki et al. [33] use nonverbal sound primarily as a state indicator, San Martin et al. [34] deploy nonverbal sound as a continuous hazard indicator with pulse frequency encoding proximity to the robot. This distinction directly shapes design requirements: event-based signals must be clear and timely, while continuous signals must remain tolerable over extended exposure without inducing sustained cognitive effort or annoyance. Hostettler et al. [35] find that some participants perceived the robot sounds as stressful or distracting. This may point to a broader blind spot in which unintended perceptual effects of nonverbal sounds are observed but rarely treated as a design problem. More broadly, sound is treated either as a designed element or as an uncontrolled condition.

Kacem et al. [24] and Roy et al. [25] demonstrate that robot motion and functional states can be mapped to nonverbal sound parameters, making robot behavior perceptible without requiring visual attention. This is directly relevant to I5.0 contexts where operators must monitor multiple systems simultaneously and visual channels are frequently overloaded.

Nonverbal sound in I5.0 contexts is gaining traction and receiving increasing attention from the research community. However, further studies are still needed to validate the

effectiveness of nonverbal sounds in industrial environments. This requires a systematic approach to sound design in order to ensure comparable and reproducible research.

4.2. Comparison of Nonverbal Sound in I5.0 vs. General HRI

The comparison of n_1 and n_2 confirms that nonverbal sound research remains concentrated in social and service robotics. The four-cluster structure of n_1 applies to n_2 , confirming thematic overlap, yet research and content within each cluster diverge substantially.

By applying the cluster structure of n_1 , the major themes of the general HRI literature become obscured. Social robotics, service robotics and emotion-based HRI could fit as standalone clusters but are instead subsumed.

The most pronounced difference appears in the Sonification cluster. While n_2 is dominated by emotion-based sound and social expressiveness, n_1 focuses on functional auditory communication and noise robustness. This confirms that I5.0's proximity-based collaboration creates different communication requirements than social robotics research addresses. But would emotions even be welcomed in an industrial use case? Based on a recent study [71] speech-based emotions from robots, like being friendly or saying "Thank you!", seem to be unwelcome among workers. Adding emotional character to nonverbal sounds on the other hand is tolerated by the study participants.

Nevertheless, methodologies from n_2 remain underutilized in I5.0 and represent concrete transfer potential:

- Layered auditory communication [53];
- Movement-based sonification [54];
- Semiotic sound analysis [45].

Data set n_2 reinforces the finding of n_1 that multimodal feedback consistently outperforms unimodal alternatives [55]. However, the application contexts in n_2 differ from I5.0 manufacturing environments and cannot substitute for industry-specific validation. Notably, I5.0-focused research creates the circumstance of rarely addressing sound in isolation, unlike social robotics research. This may appear contrary to the finding that multimodality outperforms unimodal communication, but the specific contribution and separate effects of individual modalities often remain obscured within multimodal setups. The Safety cluster reveals the most actionable transfer gap. n_2 provides foundational parameters on auditory detectability [15] and localization [61] that are prerequisite for safety sound design in I5.0 but absent from n_1 . In a specific trust repair context, even simple beep timing modulates trust to robots as effectively as verbal communication [65].

The findings on consequential sound perception [59,60] further suggest that unintentionally designed robot sounds can influence operator perception negatively, calling for proactive sound design in I5.0. Record set n_2 provides taxonomic foundations for which n_1 currently lacks an industrial equivalent. Urakami and Seaborn's [70] communication-studies framework and Nees and Liebman's [68] foundation of auditory alert types offer directly usable concepts for an I5.0-specific framework. The fact that no such framework has yet been developed represents the most actionable gap identified by this review. It also reflects the broader observation by Robinson [20] that the needed interdisciplinary approach of sound design for HRI is systematically underdeveloped.

4.3. Synthesized Research Gaps and Implications

The following research gaps are derived from both the gaps identified in n_1 and the transfer gaps emerging from the comparison of n_1 and n_2 .

- *Emotion-Related Sound Perception:* Research on nonverbal sounds in I5.0 neglects the investigation of emotional sound perception on nonverbal sounds. Dimen-

sions such as valence and arousal could support urgency communication in I5.0 but remain unaddressed;

- *Field Studies*: Research is predominantly conducted in laboratory settings, thereby neglecting environmental influences;
- *Industry-Specific Sound Design Frameworks*: Existing studies rely on ad hoc sound designs. The literature does not provide a dedicated design framework for nonverbal sound within I5.0 or industry in general;
- *Sonification Approaches for Spatial Awareness and Safety*: Sonification methods that enhance spatial awareness and safety perception in general HRI are not transferred or evaluated within I5.0 contexts;
- *Unimodal Auditory Investigations in I5.0*: Research predominantly utilizes sound within multimodal systems. This is reasonable but results in a lack of isolated investigations to detect specific contribution and limitations of nonverbal sound.

Collectively, these gaps indicate that nonverbal sound research in I5.0 has reached the proof-of-concept stage but lacks the methodological foundations for systematic advancement. One practical implication is the absence of an industry-specific sound design framework. Without structured design guidelines, sound choices across studies remain unjustified and incomparable. This implies the need for systematic interdisciplinary collaboration between robotics engineering and sound design.

A second implication concerns the integration of auditory modalities into multimodal feedback systems. Without isolating the contribution of nonverbal sound (e.g., through A/B testing), it remains unclear which communicative functions actually benefit from or require auditory support. Unimodal studies conducted under realistic industrial noise conditions should therefore be regarded as a prerequisite for informed multimodal system design, as masking effects and reduced auditory detectability may alter the effectiveness of nonverbal auditory communication in manufacturing environments. This is compounded by the underutilization of established sonification principles from general HRI, which have not been systematically transferred to I5.0.

4.4. Limitations of This Review

This review is subject to several limitations. The identification time frame was restricted to 2023 onwards, which, while justified by the I5.0 novelty and the Zhang and Fitter baseline [14], may have excluded relevant earlier contributions on industrial sound design or auditory HRI that predate the I5.0 framing. The eligibility and relevance criteria were developed iteratively during full-text assessment, which, despite group discussion among four experts, introduces a degree of subjectivity inherent to scoping review methodology. The mapping of n_2 articles onto the cluster structure derived from n_1 ensures comparability but may privilege industrial-domain categories and obscure application patterns unique to general HRI. Finally, the inclusion of only peer-reviewed publications may have excluded the relevant gray literature, particularly industrial white papers or technical reports addressing nonverbal sound in practice.

5. Conclusions

This scoping review provides an overview of the current state of nonverbal auditory communication in HRI within I5.0 and compares it with general HRI literature. As I5.0 calls for closer and more dynamic human–robot collaboration, auditory communication must move beyond traditional alarm-and-speech designs. Syntax S_1 targeted the intersection of nonverbal sound, HRI, and I5.0 manufacturing contexts. Syntax S_2 covered the broader nonverbal sound literature. Together, the two resulting record sets n_1 and n_2 yielded

50 articles, grouped into four clusters: Sonification, Multimodal Feedback Systems, Safety and Frameworks/Concepts.

Addressing RQ1, the I5.0-specific literature (n_1) remains limited but coherent in direction. Nonverbal sounds improve situational awareness, reduce cognitive workload and increase perceived safety and trust. Multimodality represents the dominant paradigm where sound is either a deliberately added signal or a consequential sound created from robot motion. Sonification demonstrates that state and motion can be mapped to nonverbal sound even under industrial noise conditions, while the safety cluster reveals a distinction between event-based and continuous nonverbal sounds, each carrying different design requirements. A consistent shortcoming across S1 is the lack of documented sound design processes, controlled studies under realistic noise conditions and evaluations of user perception.

With respect to RQ2, the comparison of n_1 and n_2 for RQ2 shows overlap in structure but differences in focus. The four-cluster framework applies to both record sets, but the weighting within the clusters differs. n_2 mainly addresses emotional effects of sound reflecting its roots in social robotics. n_1 focuses on practical applications and safety-related communication. This difference reflects the requirements of I5.0. Close, task-oriented collaboration with cobots and AMRs imposes constraints that are not addressed by social robotics research. Nevertheless, several sonification approaches from n_2 would be applicable to I5.0 and remain unused.

Two concrete research priorities emerge from this review. First, a systematic sound design framework for nonverbal auditory robot task communication is needed. It should operationalize the communicative requirements of I5.0 into designable sound parameters, bridging the gap between robotics engineering and sound design practice. Second, the problem of noisy industrial environments requires a dedicated methodological response to overcome the hurdle of testing under real-world conditions. Existing research approaches were not developed for noisy manufacturing contexts, and the perceptual robustness of nonverbal robot sounds under realistic industrial conditions remains an open question.

This review confirms the initial assessment: auditory communication is underused in I5.0, not because of missing potential, but due to a lack of validated evidence and systematic design methods. The identified literature, while limited in volume, provides a coherent foundation of empirical results and conceptual frameworks that future work can build upon. Nonverbal sound may fill the gap between simple industrial alarms and speech-based HRI and could be suited for fast, low-cognitive-load communication in shared human–robot workspaces. To exploit this potential, more empirical studies in real industrial environments are needed, along with closer collaboration between sound design, robotics and user experience.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/robotics15070121/s1>, Supplementary Materials S1 PRISMA ScR Checklist.pdf; Supplementary Materials S2 Syntaxes Fulllist.xlsx.

Author Contributions: T.S.: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing—Original Draft, Visualization; M.L.: Supervision, Validation, Writing—review & editing; S.W.: Supervision, Validation, Writing—review & editing; A.v.H.: Supervision, Validation; Writing—review & editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Bavarian State Ministry of Science and the Arts (Bayerisches Staatsministerium für Wissenschaft und Kunst) through the Technologietransferzentrum Nürnberg Land (H-H3361.NÜ/5/31). The funder had no role in this review. In accordance with

the PRISMA-ScR guidelines, the funding sources of the included articles were not systematically extracted or analyzed.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: During the preparation of this manuscript, the author used Claude.AI, Model: Sonnet 4.6 and M365 Copilot, Model: GPT-5 for the purposes of phrasing, polishing and text editing. The author has reviewed and edited the output and takes full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. European Commission. Directorate General for Research and Innovation. *Industry 5.0: Towards a Sustainable, Human Centric and Resilient European Industry*; Publications Office: Luxembourg, 2021. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Wang, T.; Zheng, P.; Li, S.; Wang, L. Multimodal Human–Robot Interaction for Human-Centric Smart Manufacturing: A Survey. *Adv. Intell. Syst.* **2024**, *6*, 2300359. [\[CrossRef\]](#)
3. Su, H.; Qi, W.; Chen, J.; Yang, C.; Sandoval, J.; Laribi, M.A. Recent Advancements in Multimodal Human–Robot Interaction. *Front. Neurorobotics* **2023**, *17*, 1084000. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Savery, R. (Ed.) *Sound and Robotics: Speech, Non-Verbal Audio and Robotic Musicianship*, 1st ed.; A Chapman & Hall Book, CRC Press: Boca Raton, FL, USA; London, UK; New York, NY, USA, 2024.
5. Cohen, Y.; Faccio, M.; Rozenes, S. Vocal Communication Between Cobots and Humans to Enhance Productivity and Safety: Review and Discussion. *Appl. Sci.* **2025**, *15*, 726. [\[CrossRef\]](#)
6. Dhanda, M.; Rogers, B.A.; Hall, S.; Dekoninck, E.; Dhokia, V. Reviewing Human-Robot Collaboration in Manufacturing: Opportunities and Challenges in the Context of Industry 5.0. *Robot. Comput.-Integr. Manuf.* **2025**, *93*, 102937. [\[CrossRef\]](#)
7. Khan, A.A.; Thurow, K. Human–Robot Interaction in Indoor Mobile Robotics: Current State, Interaction Modalities, Applications, and Future Challenges. *Sensors* **2026**, *26*, 1840. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Dua, M.; Akanksha.; Dua, S. Noise Robust Automatic Speech Recognition: Review and Analysis. *Int. J. Speech Technol.* **2023**, *26*, 475–519. [\[CrossRef\]](#)
9. Bini, S.; Carletti, V.; Saggese, A.; Vento, M. Robust Speech Command Recognition in Challenging Industrial Environments. *Comput. Commun.* **2024**, *228*, 107938. [\[CrossRef\]](#)
10. Wickens, C.D. Multiple Resources and Performance Prediction. *Theor. Issues Ergon. Sci.* **2002**, *3*, 159–177. [\[CrossRef\]](#)
11. Broadbent, D.E. *Decision and Stress*; Academic Press: London, UK; New York, NY, USA, 1971.
12. Szalma, J.L.; Hancock, P.A. Noise Effects on Human Performance: A Meta-Analytic Synthesis. *Psychol. Bull.* **2011**, *137*, 682–707. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Hermann, T.; Hunt, A.; Neuhoﬀ, J. (Eds.) *The Sonification Handbook*; Logos Verlag: Berlin, Germany, 2011.
14. Zhang, B.J.; Fitter, N.T. Nonverbal Sound in Human-Robot Interaction: A Systematic Review. *ACM Trans. Hum.-Robot Interact.* **2023**, *12*, 4. [\[CrossRef\]](#)
15. Agrawal, S.; Wessels, M.; de Heuvel, J.; Kraus, J.; Bennewitz, M. Sound Matters: Auditory Detectability of Mobile Robots. In *Proceedings of the 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)*; IEEE: Piscataway, NJ, USA, 2024; pp. 2233–2239. [\[CrossRef\]](#)
16. DIN EN ISO 7731:2008-12; Ergonomie—Gefahrensignale für öffentliche Bereiche und Arbeitsstätten—Akustische Gefahrensignale (ISO 7731:2003). Deutsche Fassung EN ISO 7731:2008. Beuth Verlag: Berlin, Germany, 2008. [\[CrossRef\]](#)
17. DIN 33404-3:2016-04; Gefahrensignale—Akustische Gefahrensignale—Teil 3: Einheitliches Notfallsignal. DIN Media: Berlin, Germany, 2016. [\[CrossRef\]](#)
18. ISO 8201 2017-11-00 EN 2805111; Alarm Systems—Audible Emergency Evacuation Signal—Requirements. International Organization for Standardization: Geneva, Switzerland, 2017.
19. ISO 11429 1996-12-00 EN 7349288; Ergonomics-System of Auditory and Visual Danger and Information Signals. International Organization for Standardization: Geneva, Switzerland, 1996.
20. Robinson, F.; Pelikan, H.; Watanabe, K.; Damiano, L.; Bown, O.; Velonaki, M. Introduction to the Special Issue on Sound in Human-Robot Interaction. *ACM Trans. Hum.-Robot Interact.* **2023**, *12*, 45. [\[CrossRef\]](#)
21. Arksey, H.; O'Malley, L. Scoping Studies: Towards a Methodological Framework. *Int. J. Soc. Res. Methodol.* **2005**, *8*, 19–32. [\[CrossRef\]](#)

22. Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.; Horsley, T.; Weeks, L.; et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Kacem, A.; Zbiss, K.; Mohammadi, A. Transforming Motion Into Sound: A Novel Sonification Approach FOR TEAMS OF MOBILE ROBOTS . In *Proceedings of the 2024 International Symposium on Flexible Automation*; ISFA 2024; ASME: New York, NY, USA, 2024. [\[CrossRef\]](#)
25. Roy, L.; Croft, E.A.; Kulić, D. Learning to Communicate Functional States With Nonverbal Expressions for Improved Human-Robot Collaboration. *IEEE Robot. Autom. Lett.* **2024**, *9*, 5393–5400. [\[CrossRef\]](#)
26. D'Attanasio, S.; Flipo, A.; Mimault, J.; Studzinska, A. Exploring User Preferences in Multimodal Human-Robot Interaction for the Human-Centered Design of an Assistant Robot. In *Communications in Computer and Information Science*; Springer: Cham, Switzerland, 2026; Volume 2548, pp. 173–189. [\[CrossRef\]](#)
27. Huang, J.; Xu, X.; Polzer, J. Towards Instructional Collaborative Robots: From Video-Based Learning to Feedback-Adapted Instruction. *J. Manuf. Syst.* **2025**, *83*, 29–45. [\[CrossRef\]](#)
28. Dei, C.; Meregalli Falerni, M.; Cilsal, T.; Redaelli, D.F.; Lavit Nicora, M.; Chiappini, M.; Storm, F.A.; Malosio, M. Design and Testing of (A)MICO: A Multimodal Feedback System to Facilitate the Interaction between Cobot and Human Operator. *J. Multimodal User Interfaces* **2025**, *19*, 21–36. [\[CrossRef\]](#)
29. Mizanoor Rahman, S.M. Communications in a Trust-Triggered Cyber-Physical-Human System for Human-Robot Collaborative Manufacturing. In *Proceedings of the 2025 IEEE International Conference on Robotics and Technologies for Industrial Automation (ROBOTHIA)*; IEEE: Piscataway, NJ, USA, 2025; pp. 1–6. [\[CrossRef\]](#)
30. Grego, G.; Nenna, F.; Gamberini, L. Enhancing Human-Machine Interactions: A Novel Framework for AR-Based Digital Twin Systems in Industrial Environments. In *Proceedings of the 17th International Conference on Pervasive Technologies Related to Assistive Environments, Crete Greece*; ACM: New York, NY, USA, 2024; pp. 456–462. [\[CrossRef\]](#)
31. Javernik, A.; Kovič, K.; Palčič, I.; Ojsteršek, R. Audio-Visual Effects of a Collaborative Robot on Worker Efficiency. *Symmetry* **2023**, *15*, 1907. [\[CrossRef\]](#)
32. Haney, J.M.; Liang, C.J. A Literature Review on Safety Perception and Trust during Human-Robot Interaction with Autonomous Mobile Robots That Apply to Industrial Environments. *IIEE Trans. Occup. Ergon. Hum. Factors* **2024**, *12*, 6–27. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Kamezaki, M.; Wada, T.; Sugano, S. Dynamic Collaborative Workspace Based on Human Interference Estimation for Safe and Productive Human-Robot Collaboration. *IEEE Robot. Autom. Lett.* **2024**, *9*, 6568–6575. [\[CrossRef\]](#)
34. San Martin, A.; Kildal, J.; Lazkano, E. Mixed Reality Representation of Hazard Zones While Collaborating with a Robot: Sense of Control over Own Safety. *Virtual Real.* **2025**, *29*, 1107. [\[CrossRef\]](#)
35. Hostettler, D.; Mayer, S.; Albert, J.L.; Jenss, K.E.; Hildebrand, C. Real-Time Adaptive Industrial Robots: Improving Safety And Comfort In Human-Robot Collaboration. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems CHI'25, New York, NY, USA*; ACM: New York, NY, USA, 2025. [\[CrossRef\]](#)
36. Chu, D.; Yu, S.; Ling, Y.; Zhao, Y.; Zhang, J. A Game-Theoretic and Multimodal Interaction Framework for Collaborative Robots in Smart Manufacturing. *Decis. Mak. Appl. Manag. Eng.* **2025**, *8*, 36–52.
37. Schneiders, E.; Fourie, C.; Celestin, S.; Shah, J.; Jung, M. Understanding Entrainment in Human Groups: Optimising Human-Robot Collaboration from Lessons Learned during Human-Human Collaboration. In *Proceedings of the Conference on Human Factors in Computing Systems-Proceedings*; ACM: New York, NY, USA, 2024. [\[CrossRef\]](#)
38. Zhang, B.J.; Orthmann, B.; Torre, I.; Bresin, R.; Fick, J.; Leite, I.; Fitter, N.T. Hearing It Out: Guiding Robot Sound Design through Design Thinking. In *Proceedings of the 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*; IEEE: Piscataway, NJ, USA, 2023; pp. 2064–2071. [\[CrossRef\]](#)
39. Wang, X.; Shen, L.; Lee, L.H. A Systematic Review of XR-Enabled Remote Human-Robot Interaction Systems. *ACM Comput. Surv.* **2025**, *57*. [\[CrossRef\]](#)
40. Khota, A.; Cooper, E.W.; Yan, Y. Synthesis of Non-Linguistic Utterances for Sound Design Support Using a Genetic Algorithm. *Appl. Sci.* **2024**, *14*, 4572. [\[CrossRef\]](#)
41. Abboud, A.; Bruno, B. Automated Generation of Emotion-Conveying Sounds for Robots. In *Proceedings of the 2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*; IEEE: Piscataway, NJ, USA, 2025; pp. 84–93. [\[CrossRef\]](#)
42. Liu, X.; Dong, J.; Jeon, M. Robots' "Woohoo" and "Argh" Can Enhance Users' Emotional and Social Perceptions: An Exploratory Study on Non-lexical Vocalizations and Non-linguistic Sounds. *J. Hum.-Robot Interact.* **2023**, *12*, 14. [\[CrossRef\]](#)

43. Andrade Pires, G.; Tsvetcoff, R.; Ayanoglu, H.; Duarte, E. Emotive Acoustics: Sound Design in Robotic Emotion Expression. A Study on Participant Generated Sounds. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer Nature: Cham, Switzerland, 2024; Volume 14685, pp. 29–41. [\[CrossRef\]](#)
44. Wolfe, H.; Su, Y.; Wang, J. Dimensional Design of Emotive Sounds for Robots. In *Proceedings of the 2024 19th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*; IEEE: Piscataway, NJ, USA, 2024; pp. 791–799.
45. Latupeirissa, A.B.; Murdeshwar, A.; Bresin, R. Semiotic Analysis of Robot Sounds in Films: Implications for Sound Design in Social Robotics. *Int. J. Soc. Robot.* **2025**, *17*, 39–58. [\[CrossRef\]](#)
46. Latupeirissa, A.B.; Panariello, C.; Bresin, R. Probing Aesthetics Strategies for Robot Sound: Complexity and Materiality in Movement Sonification. *ACM Trans. Hum.-Robot Interact.* **2023**, *12*, 42. [\[CrossRef\]](#)
47. Naphausen, E.; Muxel, A.; Willmann, J. New Design Potentials of Non-mimetic Sonification in Human-Robot Interaction. *ACM Trans. Hum.-Robot Interact.* **2023**, *12*, 13. [\[CrossRef\]](#)
48. Cuan, C.; Fisher, E.; Okamura, A.; Engbersen, T. Music Mode: Transforming Robot Movement into Music Increases Likability and Perceived Intelligence. *J. Hum.-Robot Interact.* **2024**, *14*, 48. [\[CrossRef\]](#)
49. Mannone, M.; Seidita, V.; Chella, A. The Sound of Swarm. Auditory Description of Swarm Robotic Movements. *J. Hum.-Robot Interact.* **2023**, *12*, 53. [\[CrossRef\]](#)
50. Zhang, B.; Syed, I.; Fick, J.; Fitter, N.T. A Case Study on Robot Sound Design for a Sidewalk Delivery Robot. In *Lecture Notes in Computer Science*; Springer: Singapore, 2025; Volume 15562, pp. 298–311. [\[CrossRef\]](#)
51. Robinson, F.A.; Bown, O.; Velonaki, M. Spatially Distributed Robot Sound: A Case Study. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference DIS '23, Pittsburgh, PA, USA*; ACM: New York, NY, USA, 2023; pp. 2707–2717. [\[CrossRef\]](#)
52. Arreghini, S.; Paolillo, A.; Abbate, G.; Giusti, A. Hearing the Robot's Mind: Sonification for Explicit Feedback in Human-Robot Interaction. In *Human-Friendly Robotics 2024*; Springer Nature: Cham, Switzerland, 2025; pp. 45–57. [\[CrossRef\]](#)
53. Orthmann, B.; Leite, I.; Bresin, R.; Torre, I. Sounding Robots: Design and Evaluation of Auditory Displays for Unintentional Human-robot Interaction. *ACM Trans. Hum.-Robot Interact.* **2023**, *12*, 16. [\[CrossRef\]](#)
54. Latupeirissa, A.B.; Bresin, R. PepperOSC: Enabling Interactive Sonification of a Robot's Expressive Movement. *J. Multimodal User Interfaces* **2023**, *17*, 231–239. [\[CrossRef\]](#)
55. Attfield, R.; Croft, E.; Kulić, D. A Comparison of Audible, Visual, and Multi-Modal Communication for Multi-Robot Supervision and Situational Awareness. In *Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*; IEEE: Piscataway, NJ, USA, 2024; pp. 8802–8809. [\[CrossRef\]](#)
56. Yang, Q.; Luo, S.; Tang, Z.; Wang, Q. A Human-Factor-Driven Evaluation of Multi-Modal Feedback for Enhancing Teleoperated Microrobotic Swarm Control. In *Proceedings of the 2025 9th International Conference on Automation, Control and Robots (ICACR)*; IEEE: Piscataway, NJ, USA, 2025; pp. 67–71. [\[CrossRef\]](#)
57. Song, T.; Matinfar, S.; Schütz, L.; Li, F.; Eck, U.; Navab, N. Exploring Multisensory Feedback for Robotic Ultrasound in XR. In *Proceedings of the 2025 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*; IEEE: Piscataway, NJ, USA, 2025; pp. 678–679. [\[CrossRef\]](#)
58. Matheus, K.; Vázquez, M.; Scassellati, B. Ommie: The Design and Development of a Social Robot for Anxiety Reduction. *J. Hum.-Robot Interact.* **2025**, *14*, 2. [\[CrossRef\]](#)
59. Allen, A.; Drummond, T.; Kulić, D. Sound Judgment: Properties of Consequential Sounds Affecting Human-Perception of Robots. In *Proceedings of the 2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*; IEEE: Piscataway, NJ, USA, 2025; pp. 94–102. [\[CrossRef\]](#)
60. Allen, A.; Drummond, T.; Kulić, D. Robots Have Been Seen and Not Heard: Effects of Consequential Sounds on Human-Perception of Robots. *IEEE Robot. Autom. Lett.* **2025**, *10*, 3980–3987. [\[CrossRef\]](#)
61. Wessels, M.; De Heuvel, J.; Müller, L.; Maier, A.L.; Bennewitz, M.; Kraus, J. Auditory Localization and Assessment of Consequential Robot Sounds: A Multi-Method Study in Virtual Reality. In *Proceedings of the 2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*; IEEE: Piscataway, NJ, USA, 2025; pp. 184–191. [\[CrossRef\]](#)
62. Lee, J.Y.; Eguchi, T.; Yeoh, W.L.; Okumura, H.; Fukuda, O. Development of Smart Navigation Robot for the Visually Impaired. *Artif. Life Robot.* **2025**, *30*, 265–275. [\[CrossRef\]](#)
63. Shih, K.; Boadi-Agyemang, A.; Carter, E.; Steinfeld, A. A Multi-Method Investigation of Guide Robot Characteristics for Blind and Low-Vision Users. *J. Hum.-Robot Interact.* **2025**, *15*, 5. [\[CrossRef\]](#)
64. West, A.; Chapman, M.; Lennox, B. Sonification of Ionising Radiation Data for Robot Operators. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer Nature: Cham, Switzerland, 2023; Volume 14136, pp. 141–149. [\[CrossRef\]](#)
65. Maehigashi, A.; Tsumura, T.; Yamada, S. Impacts of Robot Beep Timings on Trust Dynamics in Human-Robot Interaction. *Int. J. Soc. Robot.* **2024**, *16*, 2205–2222. [\[CrossRef\]](#)
66. Currie, J.; Giannaccini, M.E.; Bach, P. Sonic Sleight of Hand: Sound Induces Illusory Distortions in the Perception and Prediction of Robot Action. *Int. J. Soc. Robot.* **2025**, *17*, 1957–1975. [\[CrossRef\]](#) [\[PubMed\]](#)

67. Lee, A.G.; Mikhail, R.R.; Shin, M.; Grant, I.; Eden, S.K.; Shotwell, M.S.; Schlesinger, J.J. Auditory and Visual Alarm Designs Impact Clinicians' Perceived Cognitive Workload. *J. Clin. Monit. Comput.* **2025**, *39*, 1247–1255. [[CrossRef](#)] [[PubMed](#)]
68. Nees, M.A.; Liebman, E. Auditory Icons, Earcons, Spearcons, and Speech: A Systematic Review and Meta-Analysis of Brief Audio Alerts in Human-Machine Interfaces. *Audit. Percept. Cogn.* **2023**, *6*, 300–329. [[CrossRef](#)]
69. Wei, Y.; Rocher, N.; Gupta, C.; Nguyen, M.H.; Zimmermann, R.; Ooi, W.T.; Jouffrais, C.; Nanayakkara, S. Human Robot Interaction for Blind and Low Vision People: A Systematic Literature Review. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems CHI'25, New York, NY, USA*; ACM: New York, NY, USA, 2025. [[CrossRef](#)]
70. Urakami, J.; Seaborn, K. Nonverbal Cues in Human-Robot Interaction: A Communication Studies Perspective. *ACM Trans. Hum.-Robot Interact.* **2023**, *12*, 2. [[CrossRef](#)]
71. Nielsen, S.; Jochum, E.; Li, C.; Chrysostomou, D.; Ordoñez, R. "Robot Emotions Are Not Real!": Future Factory Workers' Perceptions, Attitudes, and Experience of Collaborative Robots, Conversational AIs, and AI-Empowered, Voice-Enabled Collaborative Robots. *J. Hum.-Robot Interact.* **2026**, *15*, 2. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.