Listening Between the Lines: Exploration of Augmented Reality Supported Radio Communication for Challenging Environments

Oliver Hein[1,2,*,†], Dominik O. W. Hirschberg[1,†], Florian Alt[2,1] and Philipp A. Rauschnabel[1,3,4]

[1]*University of the Bundeswehr Munich, Werner-Heisenberg-Weg 39, 85779 Neubiberg, Germany*
[2]*LMU Munich, Frauenlobstraße 7a, 80337 Munich, Germany*
[3]*University of Bamberg, Kapuzinerstraße 16, 96047 Bamberg, Germany*
[4]*University of Nottingham Malaysia, Jalan Broga, 43500 Semenyih, Malaysia*

Abstract

Radio communication remains the primary coordination channel in emergency and military operations, yet it is prone to cognitive overload, miscommunication, and information loss in noisy, high-stress environments. We explore how Augmented Reality (AR) and AI-based speech processing can support more reliable, multimodal communication by combining audio transmission with visual text overlays and automated summaries. Based on a field observation and two pre-studies ($N = 200$ & $N = 9$), we derive initial design insights and present a concept for an AI-enhanced AR radio communication system integrating speech-to-text, mesh networking, and visual message support. We propose a user study to test our system and evaluate instruction accuracy, response time, and perceived workload under realistic stress conditions. This position paper contributes empirical insights, a system concept, and a research agenda for AR-supported radio communication in challenging environments.

Keywords

Speech Recognition, Adaptive Augmented Reality, Radio Communication, Comprehension, Context

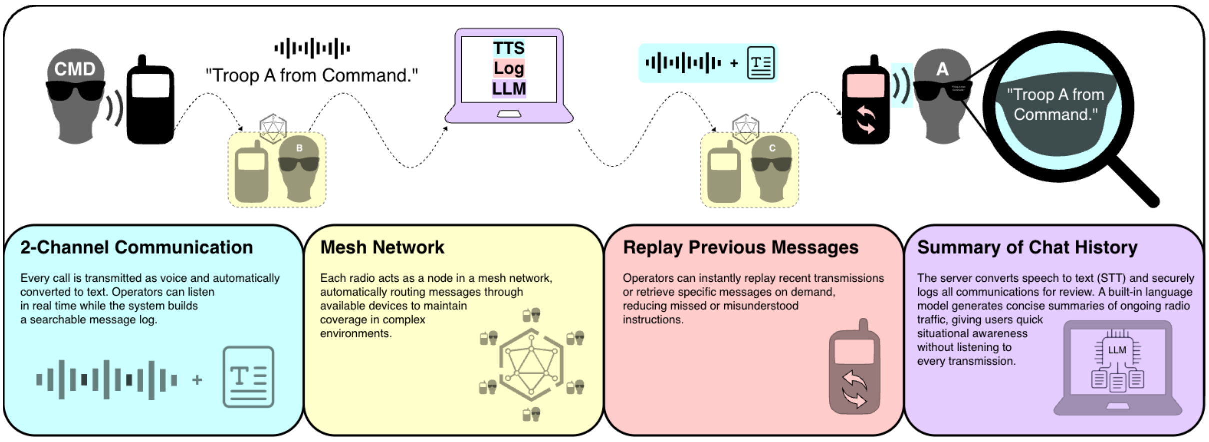

Figure 1: We present an AI-enhanced radio communication system combining voice and text transmission, mesh networking, and automated summarization. Messages are routed to a central server for speech-to-text conversion, forwarding, logging, and instant replay, while a large language model (LLM) generates concise summaries displayed on receivers' AR glasses.

CHI'26: XR4CE Workshop, April 14, 2026, Barcelona, Spain
*Corresponding author.
†These authors contributed equally.
oliver.hein@unibw.de/oliver.hein@ifi.lmu.de (O. Hein); dominik.o.w.hirschberg@gmail.com (D. O. W. Hirschberg); florian.alt@lmu.de (F. Alt); philipp.rauschnabel@gmail.com (P. A. Rauschnabel)
0009-0009-9962-0165 (O. Hein); 0009-0009-6945-8187 (D. O. W. Hirschberg); 0000-0001-8354-2195 (F. Alt); 0000-0003-2188-6747 (P. A. Rauschnabel)

1. Introduction

Events such as the ongoing Ukrainian-Russian war and the 2025 California wildfires illustrate that future crises will occur not only in cyberspace but also in the physical world. During emergency operations, fire departments, disaster control teams, and military units continue to rely heavily on radio communication. However, traditional radio communication is often insufficient in such challenging environments because spoken information is transient and can be easily missed, is susceptible to overload and distractions from background noise, may be difficult to process under time pressure, cannot be revisited later, lacks visual support for conveying complex instructions, and may suffer from transmission disruptions that obscure or distort parts of spoken messages.

To address these limitations, radio communication may be complemented with additional visual support, such as visual cues, interaction history, transcriptions of spoken sentences, and highlighted or summarized key information. Augmented Reality (AR) may be particularly beneficial in this context because it integrates virtual information into the physical environment [1, 2], making virtual information feel present in the real world [3]. Unlike conventional 2D displays or handheld interfaces, AR presents information directly in the user's field of view, enabling access to support without redirecting attention from the ongoing situation [4]. Furthermore, AR affords hands-free interaction, making it particularly relevant for communication in challenging environments, where users must simultaneously perform physical tasks and remain focused on the surroundings [5]. However, AR also requires users to allocate attention between the real world and virtual overlays, which may increase cognitive load [6, 7]. A key design challenge is selecting and combining modalities that convey information effectively without overloading users, as this directly affects AR performance in real-world settings.

Contribution Statement: In this paper, we contribute (1) design insights from two pre-studies on AR-supported multimodal instructions, (2) an innovative concept for AI-enhanced, AR-supported radio communication in challenging environments, and (3) an evaluation agenda that outlines key aspects and design challenges for future field studies.

2. Background & Related Work

Acute stress can impair working memory performance [8, 9, 10], leading to higher error rates. Task force leaders are regularly confronted with stressful, multitasking conditions that can cause cognitive overload and diminish performance [11, 12, 13]. This stress becomes a risk factor for operative forces [14]. Longer sentences also require more working memory capacity [15], so complex instructions can further increase these cognitive demands under stress.

From an information visualization perspective, visual displays support orientation, context, and dynamic feedback. Preserving interaction history allows users to retrace previous states and revisit relevant content, reducing data overload and misunderstandings [16]. According to Cognitive Load Theory (CLT) [17, 18], adapting information presentation is key to avoiding overload. AR is especially useful in challenging environments, allowing users to access information and interact hands-free without diverting attention from the real world [4, 5]. This can alleviate cognitive strain [6]. Kim et al. [19] found that although text-based AR tasks can be mentally demanding, they still yield better performance than using paper.

Research on AR in mission-critical environments has mainly focused on improving situational awareness [20, 21, 14] and providing visual instructions [22, 23, 24, 25]. Workplace-oriented projects like *wearIT@work* investigated wearable AR support for complex industrial and emergency environments [26], while ethnographic studies of emergency response practices highlight the central role of verbal coordination in dynamic situations [27]. Related work combining speech recognition and AR has largely focused on accessibility (e.g., hearing impairments [28, 29]) or language barriers [30, 31, 32]. However, no known AR solution currently logs and visually supports operational communication in real time in such environments, and design guidelines for AR interfaces supporting radio communication remain largely unexplored.

3. Preliminary Observations and Pre-Studies

To identify the challenges task forces face in demanding environments, we conducted a preliminary field observation of soldiers during a training session. The military represents a particularly relevant group, as soldiers routinely operate in complex, risky, and unpredictable situations while multitasking and making critical decisions within seconds [14]. Our observations and interviews revealed that communication and coordination were major challenges, with frequent signs of cognitive overload causing instructions to be missed or forgotten. Think-aloud feedback and follow-up brainstorming suggested that displaying conversations and key conclusions in AR could help reduce this overload. Given the lack of design guidelines and prior evaluations, we conducted two independent lab pre-studies to better understand the impact and design of displaying conversations and conclusions in AR.

3.1. Pre-Study A: Comparing Written Text Instructions in AR

Pre-study A ($N = 200$) used a between-subjects design to compare two AR text presentation formats for task instructions (continuous text vs. bullet points) in a drawing task. For each condition, a dedicated AR application was implemented on Microsoft HoloLens 2. The task was accompanied by a subjective and objective evaluation. Participants answered questionnaires about their cognitive load, affective state, feelings of fatigue, attention, and preferences after the task. Further, performance was operationalized as drawing completeness and errors.

Results Study A: Context is Key. The bullet-point format supported faster processing but was associated with a higher error rate. Continuous text improved execution accuracy but increased perceived cognitive load. Examiner ratings indicated comparable output quality across conditions. These findings suggest that instruction format should be selected based on situational requirements, particularly whether time efficiency or execution accuracy is prioritized.

3.2. Study B: The Impact of AR Text Integration on Audio Guidance

Study B ($N = 9$) [33] employed a within-subjects design to examine task performance and perceived user experience under three instruction delivery methods implemented in an AR application. The conditions were: *Audio-only* (spoken continuous text), *Audio+Text* (spoken continuous text plus the identical text displayed in AR), and *Audio+Symbols* (spoken continuous text plus a shortened symbolic representation displayed in AR). Audio and AR information were delivered via the Xreal Air 2 Pro. Participants completed three similar, complex drawing tasks, each under one of the three instruction methods. Condition order was randomized. Following each task, participants completed a questionnaire, assessing their experience, preference and feedback on the instruction methods. Performance outcomes were operationalized through error count, task completion time, and drawing accuracy.

Results Study B: AR Instruction Methods Outperform Audio-Only. Results showed, that both AR conditions reduced misunderstandings and improved perceived clarity and user experience relative to the *Audio-only* condition, reflected in lower error counts. Task completion time and drawing accuracy did not differ significantly across conditions. Participants evaluated *Audio+Text* and *Audio+Symbols* similarly overall, with individual preferences shaping which format was perceived as more usable or enjoyable. Collectively, these findings suggest that augmenting audio guidance with visual overlays can improve usability and user satisfaction in AR-assisted instruction without increasing task duration.

The pre-studies used simplified tasks. Nevertheless, they reveal key communication mechanisms, including multimodal instruction delivery and information structuring. These findings guide the design of the proposed AR-supported radio communication concept.

4. Proposed Concept & Study

Our proposed system (see Figure 1) explores an AI-enhanced approach to mission-critical radio communication, aiming to reduce cognitive load, improve message reliability, and support situational awareness. Each operator wears AR glasses, such as XREAL glasses[1], integrated into standard safety eyewear and connected to an Android smartphone. The phone serves as the main communication device, handling radio functions, mesh networking, and the local interface displayed on the glasses, while the glasses provide microphone and speaker support.

At its core, a laptop serves as the central server, which receives transmissions from the mesh network, transcribes speech to text, forwards both audio and text, logs messages, and enables instant replay of recent and past communication. It can also generate short summaries of ongoing radio traffic using a local LLM. Incoming transmissions are played as audio through the operator's speakers and shown as text on the AR display. This supports dual-channel information intake, combining visual and auditory cues. The Android phones form a digital mesh network, where each device can relay messages and help maintain connectivity in complex or degraded environments. Local buffering on the phone allows short-term storage of audio and text if the connection to the server is interrupted. These features may help reduce misunderstandings and information loss, especially in noisy or stressful situations.

Together, this hardware–software setup forms a lightweight, mobile system using off-the-shelf AR glasses, smartphones, and a portable laptop server. The approach may be applicable in domains such as firefighting, disaster response, and military operations, where clear, timely communication is critical.

Planned Study: To evaluate the concept, we plan a controlled user study examining whether AR-supported radio communication improves instruction accuracy, response time, workload, and resilience under challenging conditions. Using a between-subject design, participants will complete identical tasks across two conditions: baseline audio-only radio, and AR with live transcription. An optional third condition may include additional tools such as message replay and short message history. The study will use communication- and command-intensive scenarios such as search-and-rescue or incident response, incorporating time-critical instructions, ambiguity, interruptions, and noise to simulate realistic stressors. Participants will ideally be domain professionals working in small teams.

5. Key Challenges & Next Steps

Technical Challenges: A major challenge is ensuring reliable, low-latency speech-to-text performance in noisy, high-stress environments where accents, jargon, and radio interference are common. Hands-free interaction is essential in critical tasks, but gesture, voice, and physical controls each have limitations in harsh environments. Future work should explore robust multimodal input and evaluate its usability with gloves, protective gear, or under physical stress.

Conceptual Challenges: Operators must trust the system, especially when automated summaries or transcriptions are used, as errors could have serious consequences. Future work will focus on transparent feedback, confidence indicators, and fallback strategies. AR overlays must also be carefully designed to avoid distraction or obscuring critical information, balancing density, timing, and placement in the field of view.

Methodical Challenges & Next Steps: A key methodological challenge is designing studys that reflect real-world time pressure, cognitive demands, and environmental constraints while maintaining experimental control. Early phases will therefore use structured lab studies to assess cognitive load, attention, and interaction efficiency, followed by in-situ trials comparing AR-assisted communication with current radio practices in more realistic settings.

In parallel, we will refine the prototypes from our pre-studies to match our concept and will further iteratively refine our concept and system based on insights from each evaluation phase, with the goal of improving the system's reliability, usability, and overall impact on real-world communication.

[1]https://www.xreal.com/, last accessed March 15, 2026

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] I. Viola, M. Ahsan, O. Chatzifoti, A. Yonkov, E. Oikonomou, I. Radin, P. Mąka, A. Issam, P. Cesar, Enhancing the audience experience for vr and ar theatre with ai-generated subtitles, in: Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology, 2025, pp. 1–11. doi:10.1145/3756884.3766057.

[2] P. A. Rauschnabel, R. Felix, C. Hinsch, H. Shahab, F. Alt, What is xr? towards a framework for augmented and virtual reality, Computers in Human Behavior 133 (2022) 107289. doi:10.1016/j.chb.2022.107289.

[3] K. E. Schein, P. A. Rauschnabel, S. Praxmarer-Carus, B. J. Babin, Unpacking augmentation quality and local presence: Factors that drive effective augmented reality marketing, Journal of the Academy of Marketing Science (2025) 1–21. doi:10.1007/s11747-025-01108-2.

[4] K. P. Houweling, S. C. Mallam, K. van de Merwe, K. Nordby, The effects of augmented reality on operator situation awareness and head-down time, Applied Ergonomics 116 (2024) 104213. doi:10.1016/j.apergo.2023.104213.

[5] K. Zhang, B. R. Cochran, R. Chen, L. Hartung, B. Sprecher, R. Tredinnick, K. Ponto, S. Banerjee, Y. Zhao, Exploring the design space of optical see-through ar head-mounted displays to support first responders in the field, in: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 2024, pp. 1–19. doi:10.1145/3613904.3642195.

[6] J. Buchner, K. Buntins, M. Kerres, The impact of augmented reality on cognitive load and performance: A systematic review, Journal of Computer Assisted Learning 38 (2022) 285–303. doi:10.1111/jcal.12617.

[7] J. Baumeister, S. Y. Ssin, N. A. ElSayed, J. Dorrian, D. P. Webb, J. A. Walsh, T. M. Simon, A. Irlitti, R. T. Smith, M. Kohler, et al., Cognitive cost of using augmented reality displays, IEEE Transactions on Visualization and Computer Graphics 23 (2017) 2378–2388. doi:10.1109/TVCG.2017.2735098.

[8] S. J. Lupien, C. J. Gillin, R. L. Hauger, Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: A dose–response study in humans., Behavioral neuroscience 113 (1999) 420–430. doi:10.1037/0735-7044.113.3.420.

[9] N. Y. Oei, W. T. Everaerd, B. M. Elzinga, S. van Well, B. Bermond, Psychosocial stress impairs working memory at high loads: an association with cortisol levels and memory retrieval, Stress 9 (2006) 133–141. doi:10.1080/10253890600965773.

[10] D. Schoofs, D. Preuß, O. T. Wolf, Psychosocial stress induces working memory impairments in an n-back paradigm, Psychoneuroendocrinology 33 (2008) 643–653. doi:10.1016/j.psyneuen.2008.02.004.

[11] R. M. Shiffrin, W. Schneider, Controlled and automatic human information processing: Ii. perceptual learning, automatic attending and a general theory., Psychological review 84 (1977) 127–190. doi:10.1037/0033-295X.84.2.127.

[12] M. M. Weightman, K. L. McCulloch, M. V. Radomski, M. Finkelstein, A. S. Cecchini, L. F. Davidson, K. J. Heaton, L. B. Smith, M. R. Scherer, Further development of the assessment of military multitasking performance: iterative reliability testing, PLoS One 12 (2017) e0169104. doi:10.1371/journal.pone.0169104.

[13] M. van Beurden, L. Roijendijk, Towards a mixed reality platform for applied cognitive load evaluation, in: Engineering Psychology and Cognitive Ergonomics (EPCE 2019), Springer, 2019, pp. 123–136. doi:10.1007/978-3-030-22507-0_10.

[14] L. Chérif, V. Wood, A. Marois, K. Labonté, F. Vachon, Multitasking in the military: Cognitive

consequences and potential solutions, Applied Cognitive Psychology 32 (2018) 429–439. doi:10.1002/acp.3415.
[15] B. Sigurd, M. Eeg-Olofsson, J. Van Weijer, Word length, sentence length and frequency–zipf revisited, Studia Linguistica 58 (2004) 37–52. doi:10.1111/j.0039-3193.2004.00109.x.
[16] B. Shneiderman, The eyes have it: a task by data type taxonomy for information visualizations, in: Proceedings 1996 IEEE Symposium on Visual Languages, 1996, pp. 336–343. doi:10.1109/VL.1996.545307.
[17] J. Sweller, Cognitive load theory, in: Psychology of Learning and Motivation, volume 55, Elsevier, 2011, pp. 37–76. doi:10.1016/B978-0-12-387691-1.00002-8.
[18] J. Sweller, J. J. Van Merrienboer, F. Paas, Cognitive architecture and instructional design: 20 years later, Educational Psychology Review 31 (2019) 261–292. doi:10.1007/s10648-019-09465-5.
[19] S. Kim, M. A. Nussbaum, J. L. Gabbard, Influences of augmented reality head-worn display type and user interface design on performance and usability in simulated warehouse order picking, Applied Ergonomics 74 (2019) 186–193. doi:10.1016/j.apergo.2018.08.026.
[20] S. Julier, Y. Baillot, M. Lanzagorta, D. Brown, L. Rosenblum, Bars: Battlefield augmented reality system, in: NATO Symposium on Information Processing Techniques for Military Systems, 2000. URL: https://apps.dtic.mil/sti/tr/pdf/ADP010892.pdf.
[21] M. A. Livingston, L. J. Rosenblum, D. G. Brown, G. S. Schmidt, S. J. Julier, Y. Baillot, J. E. Swan, Z. Ai, P. Maassel, Military applications of augmented reality, Handbook of augmented reality (2011) 671–706. doi:10.1007/978-1-4614-0064-6_31.
[22] F. Ababsa, Augmented reality application in manufacturing industry: Maintenance and non-destructive testing (ndt) use cases, in: Augmented Reality, Virtual Reality, and Computer Graphics: 7th International Conference, AVR 2020, Lecce, Italy, September 7–10, 2020, Proceedings, Part II 7, Springer, 2020, pp. 333–344. doi:10.1007/978-3-030-58468-9_24.
[23] V. Kretschmer, S. Klöcker, B. Wolfgarten, R. Berner, Datenbrillen erobern die logistik: Überprüfung von augmented reality-gestützter kommissionierung in der praxis, Dokumentation des 66. Frühjahrskongresses der Gesellschaft für Arbeitswissenschaft e. V.(GfA), Digitaler Wandel, digitale Arbeit, digitaler Mensch? GfA (2020) 16–18. URL: https://gfa2020.gesellschaft-fuer-arbeitswissenschaft.de/inhalt/B.3.3.pdf.
[24] M. V. Skljarov, O. I. Sapovalov, P. V. Cernenko, S. V. Semencenko, A. A. Kaskanov, V. A. Kaskanov, The use of augmented reality technology in the training of specialists in the operation and repair of multi-purpose armored military vehicles, Visnik masinobuduvannja ta transportu 17 (2023) 153–162. doi:10.31649/2413-4503-2023-17-1-153-162.
[25] Y. Zhu, N. Li, Virtual and augmented reality technologies for emergency management in the built environments: A state-of-the-art review, Journal of Safety Science and Resilience 2 (2021) 1–10. doi:10.1016/j.jnlssr.2020.11.004.
[26] P. Lukowicz, A. Timm-Giel, M. Lawo, O. Herzog, Wearit@work: Toward real-world industrial wearable computing, IEEE Pervasive Computing 6 (2007) 8–13. doi:10.1109/MPRV.2007.89.
[27] S. Denef, L. Ramirez, T. Dyrks, G. Stevens, Handy navigation in ever-changing spaces: An ethnographic study of firefighting practices, in: Proceedings of the 7th ACM conference on Designing interactive systems, 2008, pp. 184–192. doi:10.1145/1394445.1394465.
[28] M. Mirzaei, M. R. Mirzaei, S. Ghorshi, S. Ghorshi, M. Mortazavi, M. Mortazavi, Using augmented reality and automatic speech recognition techniques to help deaf and hard of hearing people, Virtual Reality International Conference (2012). doi:10.1145/2331714.2331720.
[29] D. Watanabe, Y. Takeuchi, T. Matsumoto, H. Kudo, N. Ohnishi, Communication support system of smart glasses for the hearing impaired (2018) 225–232. doi:10.1007/978-3-319-94277-3_37.
[30] D. Wibowo, I. K. Putri, L. Saputri, Integration of augmented reality and voice recognition in learning english for children, Journal of Applied Intelligent System (2022). doi:10.33633/jais.v7i2.6119.
[31] S.-C. Tsai, S.-C. Tsai, Learning with mobile augmented reality- and automatic speech recognition-based materials for english listening and speaking skills: Effectiveness and perceptions of non-english major english as a foreign language students, Journal of Educational Computing Research (2022). doi:10.1177/07356331221111203.

[32] C. S. Che Dalim, M. S. Sunar, A. Dey, M. Billinghurst, Using augmented reality with speech input for non-native children's language learning, International Journal of Human-Computer Studies 134 (2020) 44–64. URL: https://www.sciencedirect.com/science/article/pii/S1071581918303161. doi:https://doi.org/10.1016/j.ijhcs.2019.10.002.
[33] O. Hein, S. Gökçel, F. Alt, From ears to eyes: The impact of augmented text integration on audio guidance, in: IFIP Conference on Human-Computer Interaction, Springer, 2025, pp. 510–519. doi:10.1007/978-3-032-04999-5_30.