The Auditory Contract: Designing Trustworthy Sound in IoT

Preface

Too often, sound design for smart devices is treated as a definitive, fixed layer. It is something users must accept if they wish to use the device to its full capability. There are too many variables for any single set of sounds to work universally. Hearing differences, emotional states, environmental noise, and cultural expectations all play a role. No designer can anticipate them all. If both sides, designer and listener, understand these constraints, they can become opportunities. A way to meet each other halfway. What we might call an *auditory contract* is not a prescription, but a quiet agreement to navigate meaning, trust, and attention together.

Abstract

This document introduces the concept of the *auditory contract*, the unspoken agreement between users and smart devices to trust and interpret sound as meaningful. It presents ten design guidelines for creating trustworthy, inclusive, and context-aware sound cues in the Internet of Things (IoT), supported by real-world examples in Appendix A. Designed for practitioners, researchers, and developers, it bridges ethical design with practical action.

We talk a lot about trust in design. Not enough about sound.

Every day, we accept stylised, abstract, and often artificial sounds as meaningful. From the gentle chirp of a thermostat to the bloop of a wearable or the chime of a smart assistant. These sounds are not naturalistic. They are not always intuitive. Yet we learn to interpret them, rely on them, and sometimes even feel reassured by them.

There is a quiet understanding at work here. An auditory contract. A tacit agreement between listener and device that says, *“You may not sound real, but I will trust you.”*

In the world of IoT, that trust is increasingly difficult to earn. Smart devices do not live in isolation. They coexist in the same rooms, use the same speakers, and often talk over each other. Each has its own sonic vocabulary. Few share a design language. Very few consider the combined effect of overlapping notifications, ambiguous alerts, or mood-setting tones that compete for space in our homes and pockets.

Designing for the auditory contract means thinking beyond the device. It means respecting the shared soundscape and crafting cues that are legible, ethical, and emotionally appropriate. Even when heard from the next room.

At the time of writing, these principles feel timely. That may change. In fact, it must. Sound is a temporal medium. So is trust.

Ten Guidelines for Ethical Sound in IoT

1. **Build a coherent sonic language**
   Use consistent tone, pacing, and texture to establish identity and legibility. Audit the existing soundscape. What tones, alerts, and voices already populate the environment? Consider your own device's sonic heritage. How does it relate to other products by the same manufacturer? What do users expect from its form factor or function? Very few devices are truly novel. Most extend behaviours from older tools, interfaces, or human interactions. Your sonic design is part of that lineage, whether you honour it, reinterpret it, or step away from it deliberately.
2. **Respect all listeners' attention**
   Design cues that are brief, recognisable, and unobtrusive in acoustically crowded environments. Many users will hear your device without choosing to. Others will hear it alongside competing sounds. Even Deaf and hard-of-hearing users are part of the environment and may feel effects through vibration, shared space, or group context. Design for presence without demanding attention.
3. **Use shared metaphors thoughtfully**
   Rising versus falling. Smooth versus spiky. Bright versus dull. Clear versus muffled. Fast versus slow. These metaphors can reinforce meaning, but they can also confuse. A sound that implies urgency to one user may feel aggressive to another. Spiky, jarring sounds might suggest errors or collisions, while soft, smooth tones may suggest safety or permission. Cultural and contextual norms shape these perceptions. Test them in the environments in which they will be heard.
4. **Design for learnability over time**
   Start simple. Build complexity through repetition and variation. Use sonic scaffolding. Establish a pattern before layering nuance. Cleverness is rarely helpful. People will be learning while distracted or stressed. Focus on making cues memorable, repeatable, and forgiving.
5. **Match emotion to function**
   Urgent messages should not sound casual. Routine updates should not feel catastrophic. Your user may be angry, grieving, exhausted, or overstimulated. You do not know. Choose tones that communicate clearly without assuming the listener's mood. Avoid overblown gestures. Most of the time, nothing is life changing.
6. **Support redundancy across senses**
   Not everyone hears in the same way. Use haptic and visual cues to complement sound. Also consider the physical sounds your device makes, including clicks, whirs, fans, and impact noises. They are part of the interface, even if unintended. Make them work with your design, not against it.
7. **Coexist in the acoustic commons**
   Your device is not alone. Do not dominate the space. Where possible, use the microphone to monitor ambient sound pressure and spectral balance. Wait for natural gaps before playing cues. Let others speak. Also remember to give priority. A fire alarm or smoke detector is more important than a kettle click or microwave chime. Sometimes, the right design choice is to remain silent. Listening also means recognising when your own system is causing acoustic

disruption. For example, if a sound triggers a motor or fan that blares loudly, the user may associate your cue with discomfort or confusion. That too is feedback.

8. **Consider cultural and contextual norms**
   Sounds that signal efficiency in one culture may signal rudeness in another. What works in a quiet home might irritate in a café. Some cultures experience failure cues as shaming or socially awkward. Synthetic voices may sound polite in one region but abrupt or infantilising elsewhere. Even the sound of a button press can carry different meanings depending on local design traditions. Test sounds in multiple places and with diverse users.
9. **Allow recovery, repair, and adaptation**
   Let users undo mistakes. Make feedback repeatable. Let the system learn from what users ignore, dismiss, or misunderstand. Feedback includes repeated user actions, cancelled commands, adjustments to volume, or delays in response. If everyone turns the volume down after a certain tone, that is a signal. The system should adapt to it. Listening also includes paying attention to patterns in context and time. Adaptive cues that evolve through feedback can strengthen trust over time and help people feel understood, not just instructed.
10. **Test with people, not just AI**
    Lab conditions are tidy. Real life is not. Kitchens echo. Cafés distract. Hearing does not equal listening. Cognitive load, divided attention, and competing obligations shape how sound is processed. Your cue may be perfectly audible in your office. But invisible in the wild. Always test with real users, in real-world settings, under real cognitive conditions.

Appendix A: Examples

Grouped by theme. Each entry pairs a design goal with a scenario, often including how context reshaped interpretation.

1. Success tone in office versus crash tone near a printer
2. Rising tone clear in lab versus lost in café music
3. Upbeat chime interpreted as funeral-related in another culture
4. Confirmation beep helpful to one user versus irritating to another
5. Low click for privacy respected in home versus unheard outdoors
6. Soft spoken alert trusted by some versus ignored as background noise
7. Bright ascending cue as progress marker versus mockery in some contexts
8. Long tone helpful in morning versus overwhelming at night
9. Calm voice for battery warning trusted at home versus mistaken for conversation in public
10. Beep with vibration supportive for one user versus embarrassing on public transport
11. Naturalistic fan cue reassuring to some versus triggering sensory discomfort in others
12. Playful melody joyful in personal space versus disruptive in shared work settings

These examples highlight how perception of sound is relational. It depends not only on what is heard, but when, where, and by whom.

Glossary

Soundscape: The acoustic environment as perceived by a person, including all surrounding sounds
Sonic scaffolding: The process of using simple sound patterns as a base that can be expanded as users become more familiar
Auditory masking: When one sound is difficult to hear because of interference from another sound with similar frequency or loudness

Further Exploration

Designers are encouraged to conduct real-world trials, including:

- Contextual testing in shared and private spaces
- Diary studies where users log their reactions to sounds
- A/B testing of tone variants for learnability, clarity, and annoyance

Designers may also wish to use stimulated recall retrospective verbalisation methods, where users watch or listen to a recording of their interaction and describe their thoughts and reactions. This helps capture the cognitive and emotional processes that occur in real-world use without interrupting task flow.

Future work might explore how auditory contracts shift in shared versus individual device use, or how different communities develop trust in machine-made sound.

Coda

These ten guidelines are not commandments. They are an attempt to notice things. Then to act accordingly. In five years, some of them will likely be outdated. That is not failure. It is the nature of interaction design.

While this sits apart from Chion's work on cinematic sound or Murch's thoughts on editing, I see this more as a continuation than a departure. The same instincts of timing, tone, and care also apply beyond the screen.

If you use or adapt these ideas, please cite the full publication:
Zenodo DOI: 10.5281/zenodo.15581327

Postscript

Sound is never just a signal. It is an invitation. And, at its best, it is an agreement we should not have to speak aloud. Thank you for listening. Literally, and figuratively.