Appendix A: The complex acoustic environment subjective questionnaire

The questionnaire that was used in the present study is divided into three parts. The subjects were asked to answer each part in the provided order, but within each part no restrictions were imposed on the order of their answers.

Part A: Descriptive Questions

1. Circle all types of sound from the list below, which you could hear in the scene. (Not all sounds appear in the scene.)

   a. People talking: 1 person / 2 people / 3 people / more than 3 people
   b. Other human sounds: footsteps / laughs / coughs / shouts / whisper / other
   c. Animals: birds / dogs / cats / other
   d. Television / radio playing
   e. Music playback / live
   f. Furniture dragging
   g. Paper / book rustling
   h. Computer sounds / keyboard typing / mouse clicking
   i. Kitchen-related sounds: dishes / cutlery / coffee machine / fridge / other
   j. Water flowing / drink pouring
   k. Doors and windows opening / closing / slamming
   l. Ventilation / air-conditioning
   m. Rain / wind noise / leaves rustling
   n. Electric machinery sounds: drill / compressor / saw / other
   o. Electronic gadget sounds: phones rings and beeps / alarms / other
   p. Vehicle: cars / train / airplanes / buses / trucks / motorcycle / boat
q. Specific vehicle sounds: moving / braking / honking / other

r. Any other sounds:

Where possible, note which sounds are the most dominant with an asterisk (*).

2. Does the situation take place outdoors, indoors, or in a space that combines elements from both (for example, a roofed market)? Indoors / Outdoors / Combination indoors and outdoor

3. Can you identify the type of scenario? If so, what is it?

4. How difficult were the above questions for you to answer? Extremely Easy (0) – Extremely Difficult (10)

Part B: Subjective Ratings

Imagine you are situated in [SCENE NAME] whose sound environment is being virtually reproduced now. In the following questions you will be asked to subjectively rate various aspects of this sound environment. There are no wrong or right answers.

5. How often do you experience this kind of environment? Never (0) – All the time (10)

6. How realistic do you find this audio environment? Completely Artificial (0) – Completely Realistic (10)

7. How busy is the scene? Completely Still (0) – Extremely Busy (10)

8. How loud is this scene? Very soft (1) / Soft (2) / Comfortable but slightly soft (3) / Comfortable (4) / Comfortable but slightly loud (5) / Loud but O.K. (6) / Uncomfortably loud (7)

9. Imagine you are trying to have a conversation on the phone here. How hard would it be to listen to somebody talking over the phone? Very easy (0) – Impossible (10)
10. Would you consider it annoying to sit in this environment for a long time? Are there any particular sounds you find annoying? *Not at all (0) – Extremely annoying (10)*

If so, which one(s)?

11. How unvarying / varied is the scene? An unvarying scene has no changes and is completely predictable from one moment to the next. In contrast, a varied scene repeatedly changes and sounds unpredictable or surprising from one moment to the next. *Completely unvarying (0) – Extremely varied (10)*

12. Are there any persistent / recurrent sounds that could be considered distracting? If so, indicate their distraction level. *Slightly distracting (0) – Extremely distracting (10) / Not applicable*

What particular sound(s) do you find distracting?

13. Do sounds appear spatially hard-to-distinguish or distinct? Hard-to-distinguish may imply that sounds are difficult to identify among other sounds, whereas distinct sounds are clearly separable from other sounds in the scene. *Impossible to distinguish (0) – Extremely distinct (10)*

14. Do you feel surrounded by sound? *From one direction (0) – Completely surrounded (10)*

15. Do you feel that the space is small and constrained, or very spacious? *Very small place (0) – Extremely Spacious (10)*

16. How echoic (reverberant) do you find this space? *Not echoic at all (0) – Extremely echoey (10)*

17. Would you expect that this environment be tiring for you to listen to over time? *Not tiring at all (0) – Extremely tiring (10)*

18. How demanding is focusing on different sounds in this scene? *Extremely easy to focus (0) – Impossible to focus (10)*
19. How demanding would it be for you to attend to other activities in this environment? For example, how demanding would it be to read a book? 

Extremely easy to attend (0) – Impossible to attend (10)

20. How emotionally stressful do you find this listening environment? 

Absolutely relaxing (0) – Extremely stressful (10)

21. How complex is this scene? 

Extremely simple (0) – Extremely complex (10)

22. How pleasant do you find this environment to listen to? 

Extremely pleasant (0) – Extremely unpleasant (10)

23. How much listening effort does it take you to answer the questions for this scene? 

No effort (0) – Extreme effort (10)

---

Part C: Subjective Ratings with Speech

The instructions for this part of the questionnaire are the same as the previous part. However, the following questions should be answered while focusing on the speech that is coming from the front.

24. How much does the background noise disturb the understanding of the speech in front of you? 

Not disturbing at all (0) – Extremely disturbing (10)

25. Are there any persistent / recurrent sounds that could be considered distracting to attending the speech? If so, indicate their distraction level. 

Not distracting at all (0) – Extremely distracting (10) / Not applicable What particular sounds do you find distracting?

26. How much, approximately, can you understand of the speech in front of you? (Ignore the contents.) 

Nothing (0%) – Everything (100%)
27. While listening to the speech in front of you, would you consider this noise around you to be annoying? This may be the result of a single person or thing, whose sound is annoying. If so, how annoying is it? Not at all (0) – Extremely annoying (10)

Are there any particular sounds you find annoying? If so, which one(s)?

28. How easy or difficult is it to pick out the speech in front of you from the background sounds? Extremely easy to distinguish (0) – Impossible to distinguish (10)

29. How demanding is focusing on the speech from the front in this scene? Extremely easy to focus – Impossible to focus (10)

30. Would you expect that listening to speech in this environment be tiring to listen to over time? Not tiring at all (0) – Extremely tiring (10)

31. How much effort does listening to the speech take? No effort (0) – Extreme effort (10)

32. How complex is this scene while listening to the speech? Extremely simple (0) – Extremely complex (10)

33. How realistic do you find this audio environment? Completely Artificial (0) – Completely Realistic (10)

**Appendix B: The Tucker3 three-way data component analysis method**

In the Tucker3 model, the data are organized in an $X$ array, with rows corresponding to questions, columns to environments, and tubes (the depth dimension) to subjects. Any three-way data point in the array, $x_{ijk}$, can be obtained by the summation form:

$$x_{ijk} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{iq} b_{pj} c_{kr} g_{pqr} + e_{ijk} = \tilde{x}_{ijk} + e_{ijk} \quad (3)$$
where $a$, $b$, and $c$ are the respective elements of $A$, $B$ and $C$ – the component loading matrices of sizes $I \times P$, $J \times Q$ and $K \times R$, respectively. The columns of these matrices are orthonormal and span the individual mode subspaces of respective ranks $P \leq I$, $Q \leq J$, and $R \leq K$, all with reduced dimensionality compared to the original data. $g_{pqr}$ are the elements of $G$, the $P \times Q \times R$ ‘core array’, whose elements correspond to the degree of interaction (or weight) between the components from different modes. The residuals that are not accounted for in the reduced dimensionality model are encapsulated in $e_{ijk}$. $\tilde{X}$ (with elements $\tilde{x}_{ijk}$ on the rightmost side of Eq. 3) is called the structural image of $X$, and is an approximation to the original array of size $P \times Q \times R$.

It should be emphasized that the applied three-way model takes the raw data as input, rather than correlation or covariance matrices, as is typical in one-way PCA and factor analysis. Therefore, the model does not make any assumptions on the underlying statistics of the observations and errors (Kroonenberg 2008, p. 20), which was an important property for the present study, as no knowledge was available a priori for the structure of the data. Nevertheless, the raw data had to be preprocessed before modeling, which is further described in the Data Preparation section.