The effect of sound sources on soundscape appraisal

Authors' names: K.A. van den Bosch[1], T.C. Andringa[2]

Author's affiliation:

[1] Special Needs Education and Youth Care, University of Groningen. Grote Rozenstraat 38, 9712 TJ Groningen, The Netherlands. k.a.van.den.bosch@rug.nl

[2] Artificial Intelligence and Cognitive Engineering, University of Groningen. Nijenborgh 9, 9747 AG Groningen, The Netherlands. t.c.andringa@rug.nl

ABSTRACT

In this paper we explore how the perception of sound sources (like traffic, birds, and the presence of distant people) influences the appraisal of soundscapes (as calm, lively, chaotic, or boring). We have used 60 one-minute recordings, selected from 21 days (502 hours) in March and July 2010. First, participants indicated their appraisal in terms of pleasantness and eventfulness with a joystick. A second group of participants annotated sound sources and sound events. These results were merged to yield a combined annotation of which we used the appraisal at the start and stop times of the sound events. The results supported the predictions, on theoretical grounds associated with a central role of audible safety, that traffic sounds will cause people to appraise soundscapes as more chaotic, the sound of birds chirping contributes to a calmer appraisal, and sounds indicative of people at a distance mainly increase eventfulness. These findings will facilitate the development of an automatic soundscape evaluation system able to predict the average human appraisal of a soundscape. A validated standardized measurement for the quality of soundscapes yields many possibilities for soundscape design. For example, it could suggest how to adapt soundscapes for increased audible comfort and improved quality of life.

INTRODUCTION

To reduce the negative impact of noise on the wellbeing of people, we need to know which characteristics of the audible signal induce these unwanted effects. However, an extensive review of soundscape research by Payne, Davies and Adams (2009) identified a number of gaps in the knowledge of this field of research. Amongst others it was concluded that there is a lack of knowledge on soundscape cognition, perception and classification, and a need for tools that can be used for soundscape design. In this paper we analyze the relationship between soundscape appraisal, in terms of pleasantness and eventfulness, and sound sources. We will show that particular sound sources have predictable effects on appraisal. As such this study combines psychological and physical approaches to the assessment of the quality of soundscapes to enrich our understanding of soundscape cognition. This may lead to the development of automated soundscape assessment procedures. Such a standardized automated measure for the subjective quality of soundscapes yields many possibilities for soundscape design.

Core Affect

The way we describe our moods is similar to the way we describe the state of the world around us, and both affect each other. Axelsson, Nilsson, and Berglund (2010) showed that people assess soundscapes based on their pleasantness and eventfulness. In emotion theory, Russell (2003) proposes that people use a similar set of dimensions, pleasantness and activation, to describe their moods (core affect). We consider the concept of core affect very useful in understanding the intimate and bidirectional relationship between the perceiver's mood and the appraisal of the environment (Kuppens, Champagne, & Tuerlinckx, 2012; Andringa & Lanser, 2013). In particular the combined interpretation of the dimensions of core affect and the appraisal of soundscapes yields four qualitatively different types of soundscapes: Lively, Calm, Boring, and Chaotic (Andringa & Lanser, 2013, Van den Bosch et al., submitted). Figure 1 depicts these four soundscape types, which can be classified according to their relative pleasantness and eventfulness, as well as according to the complexity of action selection and the content of audible affordances.

Audible Safety

The investigation of the relation between the listener and the environment has shown that audible safety is an important aspect of sound that is first evaluated early in the processing of auditory information (Andringa & Lanser, 2013; Van den Bosch et al submitted). For example individuals experiencing a soundscape deficient of positive indicators of safety will neither feel safe nor pleasant. This illustrates the reciprocal relationship between the (sonic) environment and moods. Andringa and Lanser (2013) conclude that the quality of soundscapes is therefore best understood by how we appraise soundscapes in terms of safety and affordances. This notion receives support from research indicating that affective evaluations of sounds are more important then acoustic properties when it comes to categorization of sounds (Berglund *et al.,* 2002).

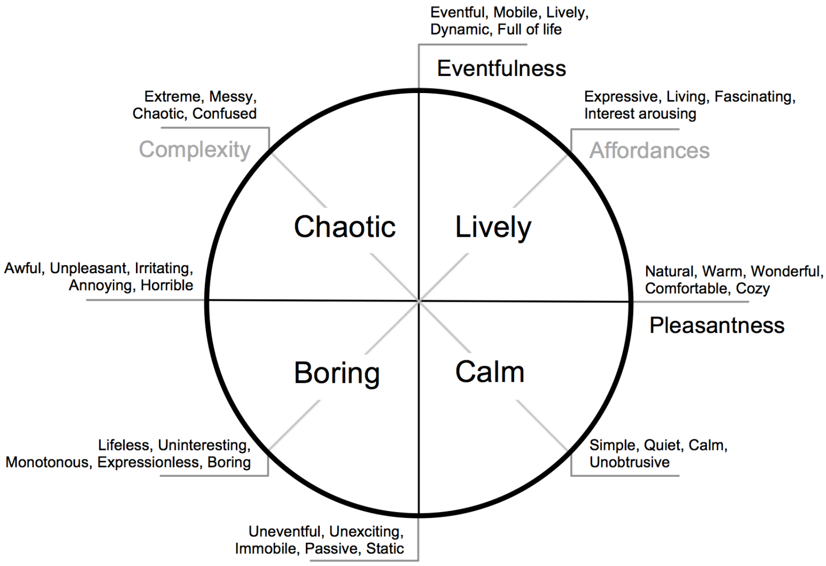

Figure 1: Four types of soundscapes (Chaotic, Lively, Boring and Calm) and their basic dimensions (Eventfulness vs. Pleasantness and Affordances vs. Complexity). In general the right side is indicative of safety, while the left side is not. (from Van den Bosch, Vlaskamp, Andringa, Post & Ruijssenaars, submitted).

The main character of a soundscape depends therefore on the prevalence and reliability of indicators of safety. Highly complex or chaotic environments are difficult to interpret (e.g., due to an overabundance of sound-producing activities) or actively indicative of insecurity. A boring environment is low on useful audible affordances and is also not indicative of safety. In contrast, a lively environment represents many affordances that provide ample interesting opportunities to attract attention and is not indicative of insecurity. The fourth environment is a calm or relaxing one because it provides ample indications of safety and allows as such full freedom to relax and recuperate (Van den Bosch, Andringa and Vlaskamp, submitted).

Expectations

Based on our theoretical framework we have a number of specific expectations that can be checked with the data that we have acquired. In particular:

1. Sound, such as from the diverse forms of traffic, mask subtle environmental sounds indicative of safety and reduce the overall sense of audible safety in addition they increase the eventfulness of the environment. The combination of the two components leads to a bias in the “chaos direction”, i.e., to the upper left.
2. Sounds, such as bird song, that connect us with a natural environment are indicative of safety and do not activate us. These sounds contribute to a calm interpretation of the sonic environment.
3. Sounds indicative of the presence of distant people do not influence pleasantness, but increase eventfulness.
4. Any sound when added to a chaotic environment will add to the chaos (possible exceptions are loud pleasant sounds, such as loud music, that may mask or distract from a chaotic “background”, but we cannot test this with our data).

METHODS

Recordings

The recordings were made on the roof of a three-story high building in the Netherlands (52°59'6.58"N - 6°32'22.79"E). The building is situated in a park-like ambiance with a secondary urban road at 50 meter and a roundabout of this road and primary urban road at 140 meters. A hotel is situated at about 100 meters across the secondary urban road. All recordings were made with a weather-proof microphone suitable for noise monitoring purposes. We used 502 hours of mono data recorded in the early spring and the summer of 2010. Of these data, 201 hours were recorded between Tuesday March 2 (2 AM) and Friday March 11 (2 PM). The remaining 301 hours were recorded in the summer from Wednesday July 21 (noon) to Tuesday August 3 (1 PM).

Selection of training and test data

Nine students[1] selected 60 70-second fragments that represented both more typical sounds and a wide variety of sounds in the recordings. The students first listened to randomly selected fragments from the available 502 hours. After half an hour of randomly listening to samples, each of the students selected ten to twelve 70-second fragments that they either found representative for the normal local soundscape or that they found were indicative of characteristic and salient events (such as passing cyclists). The students concluded that the time of the day was more indicative of the sound sources than whether the recording was made in spring or summer. On this basis, the days were split into four intervals: morning (starting at 7 am), afternoon (from 12 pm), evening (from 7 pm), and night (from 12 am). Short descriptions of the files were used to select 75 fragments, equally divided over the four intervals (also stemming from both spring and summer). This set of 75 fragments was reduced further, through removing the more redundant examples. This led to the final selection of 60 70-second fragments (15 for each part of the day).

Annotation of appraisal

The final set of 60 fragments was annotated with a joystick method (Doesburg, 2013) by the same 9 students. All students annotated all files by concurrently indicating the pleasantness and eventfulness with a joystick.

Event annotation

The Sound Recognition class of 2014 conducted the annotation of the sound sources and sound events that occurred in the fragments. This annotation was performed with a different tool (van der Linden, 2012) that allowed students to play the fragment, or part of it and inspect it with the visual aid of a cochleogram (Andringa, 2002). The annotators selected for each event an interval in the cochleogram and indicated the event class; new class labels could be introduced if necessary. The combination of the interval (start and stop times) and the selected label was stored as an event annotation. The annotators had no access to appraisal information, nor did they know their annotation would be combined with appraisal. All fragments were annotated at least twice with a maximum of five times. The results were merged to yield a combined annotation.

Data used

In this paper we combined the appraisal annotation with the event annotation to yield 922 unique annotations with on average 15 annotated events per fragment (one every 4.5 second). For each annotation we use the label and the appraisal at start and end time of the annotation. We will analyze three classes of events: traffic (N=325), birds (N=350), and people sounds (N=44). We only perform descriptive statistics.

[1] Sound Recognition course, Artificial Intelligence, University of Groningen, 2013

RESULTS

Traffic

We will first analyze the effect of the traffic class on appraisal. The traffic class consists of the labels: airplane, brake, bus, car, circuit, engine, exhaust, garbage truck, helicopter, motorcycle, scooter, traffic, train, and truck. The analysis results are depicted in Figure 2. The first thing to notice in the left panel is the rather narrow and localized distribution of the diversity of different sound classes within the appraisal plane. This validates a) the use of the average appraisal and b) the merging of a diversity of traffic events into a single class. A second thing to notice is that most arrows, each connecting appraisal at the beginning to the appraisal at the end of the annotation, are more or less parallel and in the same direction. Although the left panel may suggest a large number of arrows towards the calm quadrant, this is actually a small minority. The direction distribution in the "rose-plot" and the "compass-plot" in the upper right of Figure 2 shows that the direction towards he upper left – chaotic – quadrant is dominant. This directionality is also reflected by the change in the prevalence of the four quadrants of the appraisal field as depicted in the two panels in the lower right. The left panel of these indicates the prevalence of being in a particular appraisal quadrant at the start of the event and the panel in the lower right the prevalence of the appraisal quadrant at the end of the event. Compared to the beginning of the event the prevalence of the chaotic quadrant is much higher at the cost of the other quadrants. Overall this pattern is strongly supportive of a traffic-induced bias towards chaos, consistent with expectation 1.

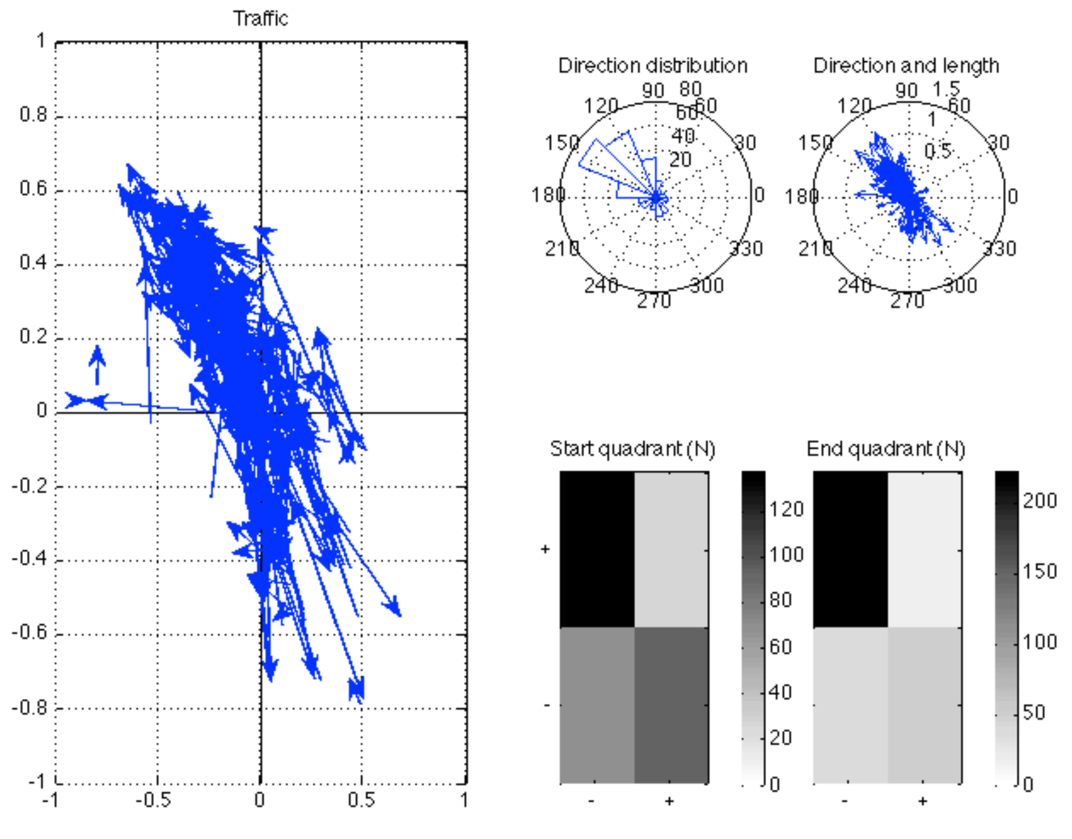

Figure 2: Analysis of the appraisal effects of traffic sounds. The direction distribution shows that most traffic corresponds consistently to an appraisal shift towards a chaotic interpretation (upper left).

Figure 3 shows a similar analysis as Figure 2, but now for the bird sounds (N=350), consisting of the labeled annotations: bird, crow, pigeon, owl, and raven. Here we see a quite different pattern in the form of two main directions: towards chaos and towards calmness, where calmness bias is stronger. In addition, prevalence matrices in the lower right show that the presence of bird sounds leads to a prevalence reduction in the boring and lively quadrants and a more prominent role of the chaos and the calm quadrant.

A more precise analysis shows the presence of two "attractors": two in the calm quadrant and one in the chaos quadrant. This offers support for expectations 2 (birds are associated with calmness) and 4 (any contribution to a chaotic environment will contribute to the chaos).

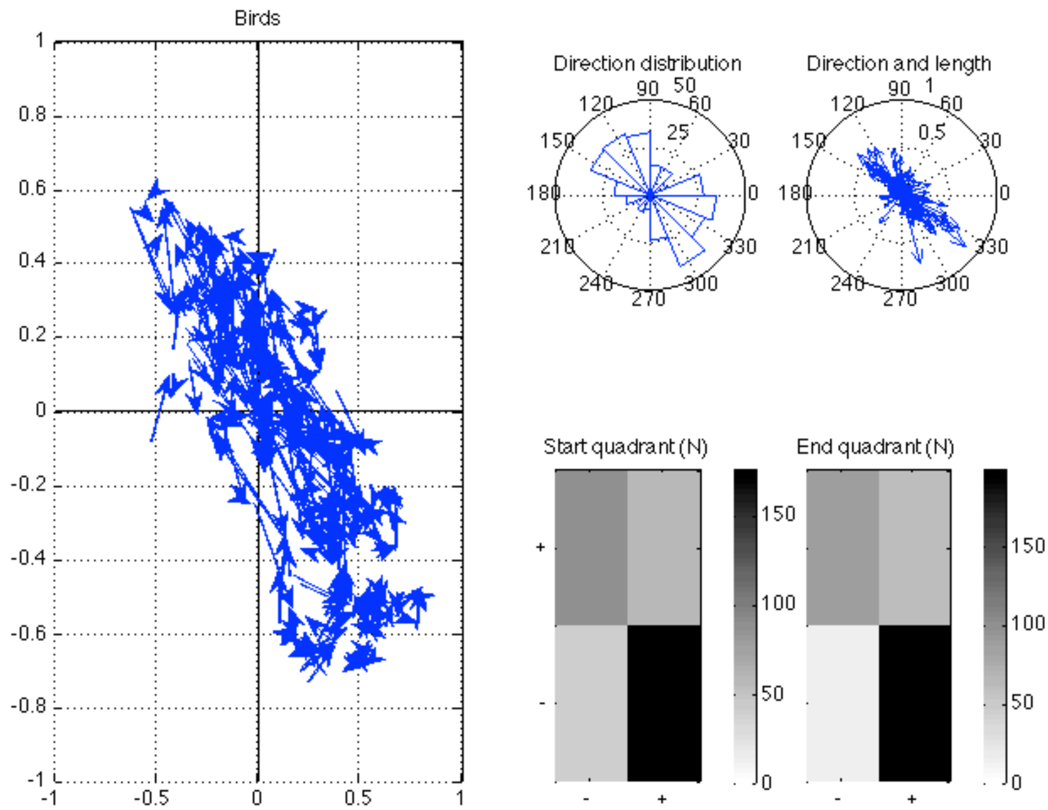

Figure 3: Analysis of the appraisal effects of bird sounds. The direction distribution shows that most bird sounds corresponds to an appraisal shift towards a calm interpretation (lower right).

Figure 4 shows the analysis of the people class that contains labels indicative of the presence of people via the labels: child, cry, footsteps, laughter, shouting, speech, and cyclist. Here we see that the perception of events indicative of people in calm situation contributes to the eventfulness (expectation 3), while the same labels in a chaotic environment add to the chaos (expectation 4). The prevalence matrices show that this class leads to an increase in eventfulness and especially a shift towards liveliness.

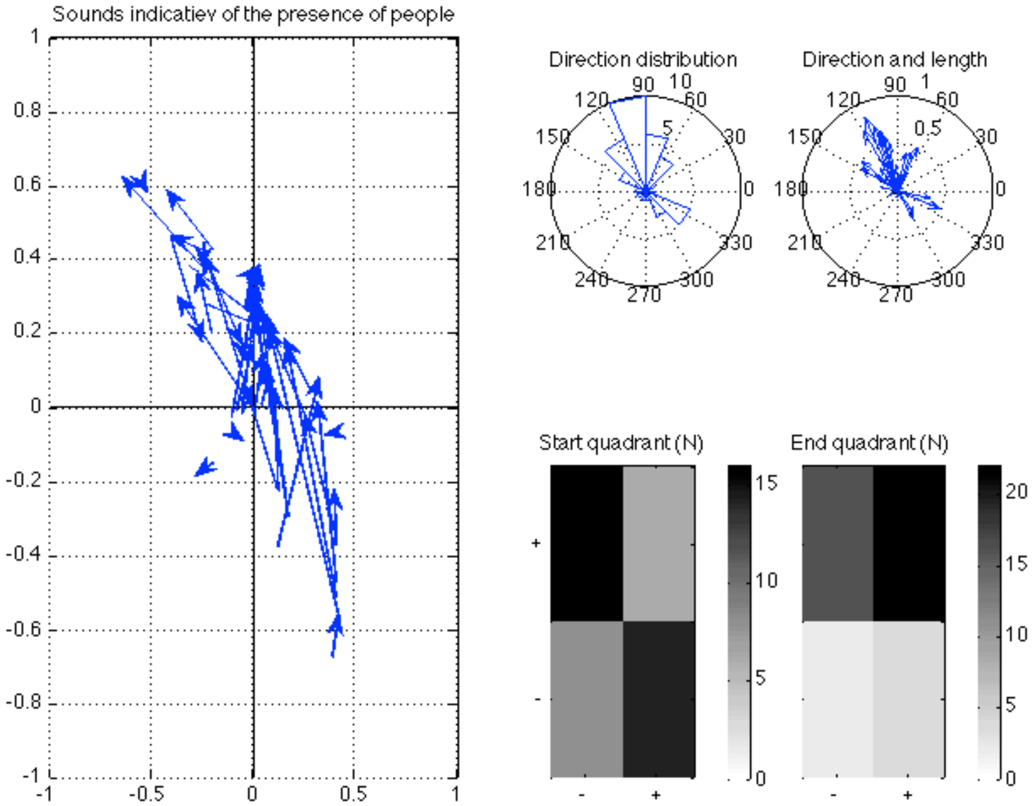

Figure 4: Analysis of the appraisal effects of sounds indicative of the presence of people. In calm situations this contributes to eventfulness and an increase in liveliness and in more chaotic situations these contribute to the “chaos”.

CONCLUSION

The results of this study are consistent with the expectations as formulated in the introduction of this paper, based on audible safety. Support was found for the prediction that traffic sounds will cause people to appraise soundscapes as more chaotic. Results also indicate that the sound of birds chirping indeed contributes to a calmer appraisal of the environment, except in chaotic soundscapes where they contribute to the chaos. Lastly, it was shown that sounds indicative of the presence of distant people do not so much influence the pleasantness, but do increase the eventfulness. The main limitation to this study is the skew distributions of data, with more data in the chaotic and calm quadrant, compared to the lively and boring quadrant. This is however a property of this specific type of city soundscape. Other (rural and urban) soundscapes should be included in further research to assess a more complete relation between event perception and changes in appraisal.

This study not only helps to understand which part of the signal is attended to and which signal components are appreciated, it also proposes important functional characteristics of soundscape appraisal. This will facilitate the development and validation of an automatic soundscape evaluation system able to predict the average human appraisal of a soundscape. A validated standardized measurement for the quality of soundscapes yields many possibilities for soundscape design. For example, it could give direct suggestions on how to adapt soundscapes for increased audible safety and improved quality of life.

REFERENCES

Andringa, T. C. (2002). *Continuity preserving signal processing* (Doctoral dissertation, PhD dissertation, University of Groningen, see http://irs. ub. rug. nl/ppn/237110156).

Andringa, T. C., & Lanser, J. J. L. (2013). How pleasant sounds promote and annoying sounds impede health: A cognitive approach. *International journal of environmental research and public health*, *10*(4), 1439-1461. doi:10.3390/ijerph10041439

Axelsson, Ö., Nilsson, M. E., & Berglund, B. (2010). A principal components model of soundscape perception. *The Journal of the Acoustical Society of America*, *128*(5), 2836-2846. doi:10.1121/1.3493436

Berglund, B., Hassmén, P. & Preis, A. (2002). Annoyance and spectral contrast are cues for similarity and preference of sounds. *Journal of Sound and Vibration*, **250** (1), 53-64.

Doesburg, I. (2013). Using a joystick to express your opinion about a sonic environment. (T.C. Andringa, Ed.). University of Groningen, Bachelor Artificial Intelligence. http://irs.ub.rug.nl/dbi/5130a0cf685fd

Kuppens, P., Champagne, D., & Tuerlinckx, F. (2012). The dynamic interplay between appraisal and core affect in daily life. *Frontiers in Psychology*. doi:10.3389/fpsyg.2012.00380/abstract

Payne, S. R., Davies, W. J., & Adams, M. D. (2009). Research into the practical and policy applications of soundscape concepts and techniques in urban areas.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*. doi:10.1037/0033-295X.110.1.145

Van den Bosch, K.A., Andringa, T.C and Vlaskamp, C., (submitted). Safe and Sound: the importance of audible safety for people with intellectual disabilities.

Van den Bosch, K.A., Vlaskamp, C., Andringa, T.C., W.J. Post and Ruijssenaars, A. (submitted). The quality of soundscapes in residential facilities for people with profound visual and intellectual disabilities.

Van der Linden, R. (2011). Development of a tool for Soundscape Annotation. What do we hear when we listen? (T.C. Andringa. Ed.). University of Groningen, Master Artificial Intelligence. http://irs.ub.rug.nl/dbi/4ef34da0110a1